ISSN-0043-7956

WORD

JOURNAL OF THE INTERNATIONAL LINGUISTIC ASSOCIATION

EDITORS

RUTH M. BREND
JOHN R. COSTELLO
EUGÉNIO CHANG-RODRÍGUEZ
SHEILA M. EMBLETON

ARTICLES

On defining linguistic periods: gradients and nuclei . . . . James Fife 1
Animateness and shape in classifiers . . . . Setsuko Kiyomi 15

SELECTED PAPERS FROM THE 36TH ILA ANNUAL CONFERENCE: INDO-EUROPEAN AND THE INDO-EUROPEANS, PART I

Indo-European linguistics today . . . . Winfred P. Lehmann 37
Sir William Jones, language families, and Indo-European . . . . Garland Cannon 49
The Nostratic macrofamily (with special reference to Indo-European) . . . . Allan R. Bomhard 61
Genealological linguistics: Progress and pitfalls . . . . Bernard H. Bichakjian 85

NOTES AND DISCUSSION

On the etymology of Portugese *SAUDADE*: an instance of multiple causation? . . . . Leo Pap 97

REVIEWS

Humphrey Tonkin and Karen Johnson-Weiner, eds. *Overcoming language barriers: The human/machine relationship* . . . . Eric Wheeler 103
Thomas M. Stephens. *Dictionary of Latin American racial and ethnic terminology* . . . . Stephen O. Murray 105
August Dauses. *Theorien des Sprachwandels: Eine kritische Ubersicht* . . . . Raimo Anttila 107
John Victor Singler, ed. *Pidgin and creole tense-mood-aspect systems* . . . . Robert B. Le Page 110
Cognitive Linguistics 1.1 and *Rivista di Linguistica* 1.1 . . . . John Hewson 114
R. Hasan & J. R. Martin, eds. *Language development: learning language, learning culture* . . . . Nigel Love 117
Nancy C. Dorian, ed. *Investigating obsolescence. Studies in language contraction* . . . . Lenora A. Timm 124
Gerald Knowles. *Patterns of spoken English: An introduction to English phonetics* . . . John Algeo 134
Robert A. Hall, Jr. *A life for language: A biographical memoir of Leonard Bloomfield* . . . . Stephen O. Murray 138
Jon P. Dayley, *Tümpisa (Panamint) Shoshone dictionary* . . . . Charles Peck 145
Jon P. Dayley, *Tümpisa (Panamint) Shoshone Grammar* . . . . Charles Peck 147
Alice Shepherd, *Wintu Texts* . . . . Charles Peck 149
James M. Crawford, *Cocopa dictionary* . . . . Charles Peck 150
William Shipley, ed. *In honor of Mary Haas: From the Haas Festival Conference on Native American linguistics* . . . . Matthew S. Dryer 151
Stig Johansson and Knut Hofland. *Frequency analysis of English vocabulary and grammar based on the LOB corpus*. . . . . John Algeo 154

VOLUME 43, NUMBER 1 APRIL, 1992

INTERNATIONAL LINGUISTIC ASSOCIATION

formerly Linguistic Circle of New York

HONORARY PRESIDENT: **André Martinet,** Sorbonne & École Pratique des Hautes Études
PRESIDENT: **Alice H. Deakins,** William Paterson College
VICE-PRESIDENT: **Hermann W. Haller,** Queens College, CUNY
RECORDING SECRETARY: **Julius Purczinsky,** Hunter College
CORRESPONDING SECRETARY: **Edward G. Fichtner,** Queens College, CUNY
TREASURER: **Theodore S. Beardsley Jr.,** The Hispanic Society of America

EXECUTIVE COMMITTEE: The preceding and

Marian F. Astuti, Long Island University
Ruth M. Brend, Michigan State University
Eugenio Chang-Rodriguez, Queens College, CUNY
John R. Costello, New York University
Sheila M. Embleton, York University
Harold Homa, West End High School, NY
Franklin Horowitz, Teachers College, Columbia University
Diana Kao, City College, CUNY
JoAnne Kleifgen, Teachers College, Columbia Univ.
Herbert Pilch, Universität, Freiburg
Julius Purczinsky, Hunter College
Solomon I. Sara, Georgetown University
Alan R. Thomas, University of North Wales
Johanna J. Woltjer, New York, N.Y.

BOARD OF EDITORS: **Ruth M. Brend** (Michigan State University) Managing Editor; **Eugenio Chang-Rodríguez** (Queens College, CUNY); **John R. Costello** (New York University); **Sheila Embleton** (York University), Review Editor.

WORD, the journal of the International Linguistic Association is published three times a year, in April, August, and December. Except as noted below, membership in the Association includes a subscription to *WORD*. As of January 1, 1985, the schedule of annual dues is as follows: regular member, $35.00; joint member (two members at the same address, receiving one copy of *WORD*), $22.50 each, or $45.00; ILA membership without *WORD* subscription, $15.00. Applications for membership, or for an institutional subscription may be sent to the Treasurer, **Dr. T.S. Beardsley, Jr.** / The Hispanic Society of America / 613 West 155 St. / New York, N.Y. 10032 / U.S.A. Payment must be made by check or money order in United States funds, payable to the International Linguistic Association. Changes of address and claims for missing issues may be sent to the Corresponding Secretary, **Dr. E.G. Fichtner** / Dept. of Germanic, Slavic and East European Languages / Queens College / CUNY / Flushing, N.Y. 11367-0904 / U.S.A.

The Board of Editors invites contributions to the journal in the form of original articles on the structure, function, or historical development of natural languages, or on theoretical questions related to these. Manuscripts (in four copies) may be sent to the Managing Editor, **Dr. R. Brend,** 3363 Burbank Dr. / Ann Arbor, MI. 48105 / U.S.A. Manuscripts must conform to the Style Sheet for *WORD*, published most recently in Volume 35, No. 1 (April 1984); a copy may also be obtained from the Managing Editor. Manuscripts of articles accepted become the property of the International Linguistic Association.

Books or journals for review or listing may be sent to the Review Editor, **Dr. Sheila Embleton** / Linguistics / York University / 4700 Keele St. / North York, Ont. M3J 1P3 / Canada. Unsolicited reviews cannot be accepted. Anyone wishing to contribute a review is asked to consult with the Review Editor before submitting it.

The Editors are jointly responsible for the contents of the Journal.

Copyright 1992 by the International Linguistic Association. All rights reserved.

JAMES FIFE

On defining linguistic periods: gradients and nuclei

Abstract. In this paper I discuss some considerations which enter into the definition of a linguistic period, i.e. a particular phase in the history of a language. I propose that historical varieties of a language be treated more on a par with synchronic variation. Our definition of a period must therefore avoid employment of arbitrary, extra-linguistic boundary-markers and should recognize and allow for graded types of categorization to supplement the usual metaphor of periods as discrete entities. What I am proposing is to bring our studies of historical variation closer to what has long been recognized as necessary in synchronic variation work. In what follows I will show how synchronic and diachronic variation have been treated differently by linguists. Then I put forward some programmatic suggestions on how to revise our definition of linguistic periods in the light of this comparison. Finally I provide a brief illustration of how this all works by considering a problem of periodization in the history of Welsh.

1.1 I start out by reviewing some of the ways in which the concept of a language variety has been described differently in the synchronic and diachronic spheres, or what I will also refer to respectively as the **latitudinal** vs. **longitudinal** views of language variation.

The branch of linguistics which has the longest history in studying latitudinal variation is dialectology. The relatively newer field of sociolinguistics shares with dialectology a concern for the latitudinal variation of language behavior, but on an axis of social position rather than on geographic distribution. Both fields provide us with similar conclusions, and my remarks refer to both in what follows, but I concentrate on the former, since it shows the clearest signs of having undergone an evolution in its treatment of variety boundaries.

Dialectology, in its widest sense, is that field of linguistic studies which attempts to define and describe the subvarieties of a single language. Francis (1983:1); Petyt (1980:29-30). When one talks about a particular variety of a language it is usually understood as some pattern of linguistic behavior which shares many features with a somehow-defined superordinate system of grammar, which we can label the **standard language,** but also has certain consistent differences from the standard. See Francis (1983:10). Of the many theoretical questions which are raised by this fairly intuitive notion of variety, one which I want to address here concerns the isolation of specific **boundary-markers** for these varieties, that is, determining the features which define the outer limits of one variety and separate it linguistically from another.

The type of boundary-marker most commonly employed in classical dialectology, as I have mentioned, is geographic distribution. Thus in classical dialectology a ''dialect'' is a variety of a standard language defined by certain geographical boundaries. Petyt (1980:27); Chambers & Trudgill (1980:7-8). Although this boundary criterion has been the mainstay of dialectology and is still commonly used (for convenience sake rather than empirical adequacy), it has long been recognized as not fully sufficient in describing the facts of latitudinal varieties of language. Petyt (1980:27-30); Chambers & Trudgill (1980: 6-8). For, while it is true that people who live close together often speak in similar ways, geographic proximity is not the sole determiner of dialect affiliation. We can see cases like English, where, for many historical and other reasons, the dialects of a small area like Great Britain show greater variation than those of a geographically dispersed area like North America. Likewise we have the cases of non-mainstream communities, such as Jews or Romani, living in close geographical contact with others but still having a pattern of speech in the same language which is different from that of the majority speakers. In the same vein, the attempts by linguists, among others, to use various other non-linguistic boundary markers to define linguistic groups proved futile. Over half a century ago the great anthropological linguist Franz Boas warned against any simple-minded equation of linguistic grouping with race or cultural type, since it is clear that all three are independently variable (Boas 1940).

Thus the notion of what marks a latitudinal linguistic boundary evolved during the history of dialectology so that now it is generally recognized that extralinguistic criteria alone are inadequate for defining synchronic varieties. Thus Russian is a language spoken by millions of non-Europeans; Spanish is spoken by peoples with radically different cultural backgrounds; English is spoken by people of differing racial stocks. Moreover, it became clear with time that it was unreasonable to attempt to find any hard and fast boundaries between varieties. See Francis (1983:158). That is to say, the boundaries between two varieties are not like a political frontier on a map, but more like the outer limits of a sphere of influence. It was found to be far too arbitrary to set any single, invariable geographic or isoglossic boundary which would say definitively here is where Plattdeutsch ends and Saxon begins. The famous case of the Rhenish Fan (Bloomfield 1933; Chambers & Trudgill 1980:106-107) is a prime example: the facts of distribution show that one variety gradually shaded into another along a number of parameters. It would be misleading and artificial to claim, for example, all speakers north of the [dat]/[das] isogloss are Low German speakers and all to the south are High German. The solution settled on by dialectologists as the most accurate reflection of the facts was to see the bundle of isoglosses as forming a transition zone between varieties, a **gradient** of linguistic behavior vis-à-vis their classification into sub-types of languages. Petyt (1980:60-61).

1.2 A related difficulty is the problem of defining levels of language variety: Is a language, as someone once claimed, merely a dialect with an army and a navy? or are there linguistic criteria? Such concepts are notoriously difficult to define in a binary system of categorization, using some set of criterial features for defining a grammatical system as one or the other. Many definitions relied heavily on extra-linguistic factors which seemed to miss the point that what constitutes a language or a dialect should be definable on a purely or largely linguistic basis. Again in variation studies, especially in sociolinguistics, such labels as 'language' or 'dialect', which in our everyday folk definitions rely on many non-linguistic factors, can best be described in terms of gradients, that is, linguistic patterns show degrees of being more like a language or more like a dialect, depending on the degree to which the pattern used is standard or divergent, historically or socially. Petyt (1980:14-15). This does not exclude extra-linguistic features from our definition, but combines them with purely linguistic ones.

1.3 Just to summarize what I have said so far about studies of latitudinal varieties, it has been found that purely non-linguistic boundary markers are inadequate. On consideration of purely linguistic factors, the distribution of criteria suggest a gradient-effect rather than rigid division into binary categories. The state of the art appears to be to recognize synchronic varieties as differing on the basis of their linguistic features, but not so discretely as to ignore the significant overlap between them.

2.0 So well-established now is this conception of latitudinal varieties that it seems like mere common-sense. Yet it is surprising to see that this same common-sense has not been so thoroughly applied to the other axis of language variation, in diachronic studies. We can see this most clearly in the criteria used to isolate historical periods in various languages.

2.1 It is commonplace for extra-linguistic boundary markers to be the primary factors in isolating linguistic periods. For instance, Old English is frequently defined in relation to the Norman Conquest, e.g. Williams (1975:65-66); Bourcier (1981:119-23); Pyles (1971:152-54). Nist (1966) treats Old English as coming to an end with the close of the *Anglo-Saxon Chronicle* in 1154. Medieval Greek is delimited by the two historical events of the founding and Latin capture of Constantinople (Lockwood 1972:6-7). The 'Golden Age' of Classical Latin is said to end with the death of Augustus (*Ibid*:24). Middle Norwegian dates from the advent of the Black Death in 1349 (*Ibid*.: 124). Classical Sanskrit begins with the codification of Panini (Whitney 1889: xiv). Classical Nahuatl is the language preserved in five early grammars (Newman 1967:179-81). Modern Welsh is said to begin with the 14th-century poetry of Dafydd ap Gwilym (Morris Jones 1913:7). Old Dutch is the language of a single sentence which pre-dates all Middle Dutch texts (Brachin 1985:9).

These and many other examples can be taken to show that the historical periods of a language are often defined on solely extra-linguistic criteria, such as manuscript tradition, the supposed influence of a single author, a convenient historical hiatus, the form given it by an influential grammarian, or because of evidential poverty. None of these necessarily have any direct relation to the *linguistic* characteristics of the material. Certainly it is often the case that different manuscript traditions show differences in the *form* of the language used, or that literary style may change after the life of an important writer. But this is using extra-linguistic criteria to define diachronic varieties of a language, analogous to defining synchronic ones by such features as race and geography.

2.2 It seems clear that historical periods of a language should be treated in the same manner as other varieties. That is, we should think of periods as 'longitudinal dialects' of the language. If that is so, we must recognize the need to abandon purely extra-linguistic criteria for defining periods, as we have done for dialects. We know that latitudinal varieties cannot be accurately defined on the basis of purely non-linguistic criteria, we should not feel justified to do the opposite with longitudinal varieties. Instead of isolating periods simply because the evidentiary tradition changes in quantity, form or style, we should recognize a new category of **attestation group,** that is, a certain tradition of attested forms of the language, whether or not these qualify as a separate period or perhaps overlap different periods. This will help keep our extra-linguistic criteria from taking over what is primarily a matter of linguistic factors.

Moreover, historical periods are often treated as monolithic in nature, if only because the boundary markers are so discrete. Because a period is sometimes defined relative to a single person's lifetime or a fixed historical event, the boundaries between one period and the next seem at times very abrupt. This makes periods marked-off and immutable in a way we would never attribute to synchronic varieties. We are prepared to recognize all manner of gradient transition between different dialects (see e.g. the arguments of Bailey 1973), but not always so ready to accept 'fuzzy' boundaries between, say, Middle and Modern English. This is somewhat exaggerated, since scholars often acknowledge that picking a date for the start or end of a phase can be arbitrary. But there is no denying that the concept of gradient varieties has not had the same wide-spread application in historical linguistics it has had in dialectology.

2.3 If, as I claim, a linguistic period is simply a longitudinal dialect (or dialects) of the language, we must reconcile our theoretical treatment of varieties on the different axes of variation and make them more compatible. We could of course make dialects more like periods, but the history of dialectology shows that this would be erroneous. It therefore appears more promising to make our concept of linguistic period more like our understanding of synchronic varieties.

3.0 I would like to turn now to some (admittedly programmatic) discussion of how we could do this and briefly illustrate this with a problem in the periodization of Welsh. I first want to give the discussion a theoretical basis by relating the phenomena described above to the concept of **prototype** as it is used in cognitive linguistics. See Lakoff (1987) for a detailed discussion of this concept and its justification.

3.1 In cognitive linguistics the usual method of arranging items into binary categories is often rejected in favor of the theory of categorization by prototype. Under the classical theory of categories, items which possess all the criterial features of a category are considered members of that category. In cognitive theory, deriving from the results of much recent research in cognitive psychology, categorization is also seen to rely crucially on the notion that items can be members of categories by degrees, that is, that some members of a category are better examples of that category than others. Some members are more central and others more peripheral, depending on the degree to which they are judged good examples of the category, but all are basically tokens of the one type. The best example (or class of examples) of a category is called the prototype, and other potential members are judged by comparison to it. In this theory of categorization, all members of a category may not have the same criterial features and there are cases where two entities are in the same category although they have no criterial features in common.

How this is accomplished takes us beyond the immediate scope of this discussion; the important point is that entities can be treated as belonging to the same category merely by being judged relatively better or worse examples of the prototype. This does not make categorization a free-for-all: formation and development of human categories is most definitely governed by discernible principles. Among those mentioned by Lakoff (1987:95-96) are centrality (i.e. categories are structured around prototypes which are more central examples of the category than more peripheral ones), chaining (i.e. peripheral examples may belong to a category through a series of transitive inferential links with more central members) and idealized cognitive modeling (i.e. that conventional logic structures can apply in determining membership). To illustrate with the example inspiring the title for Lakoff's book, the Dyirbal language utilizes four classifiers to partition the conceptual universe. One classifier organizes men and most animals, another applies to women, water, fire and dangerous things. In addition to these relatively central members of each category, the first also classifies the word for ''moon'' and the second classifies ''sun''. As Lakoff explains (based on the analysis of Dixon 1982), the moon comes under classifier I because in traditional Dyirbal myth, the moon is masculine, while the sun is feminine, as well as being naturally associated with fire. See Lakoff (1987:92-94). The sun and moon are members of their respective classes by virtue of their being linked to the two sexes by chaining idealized cognitive models of gender and mythology. Thus, while we can say in general that classifier I pertains to males and classifier II to females, there are members of these same classes which are more peripheral examples than the central prototype.

I introduce this concept here not for the purpose of arguing that actual speakers necessarily look at historical texts and consciously try to steer towards them as a standard. For one thing, a synchronic speaker obviously has only earlier prototypes available for consideration. Although I am sure there is more conscious effort at conforming to recognized linguistic prototypes than has been previously acknowledged, my intent here is simply to establish that categorization by prototype is a legitimate theoretical model *for linguists* to describe the historical lodestars which define the periods of a language.

3.2 The concept of a prototype is readily applied to the problem at hand, for we can see that the structure of a category based on prototypes accounts for the gradient effect seen in language varieties. In the case of latitudinal variation, we can define a relatively cohesive **nucleus** of linguistic behavior which serves as a prototype; in this case, the grammar of the standard language (this is not unlike Weinreich's (1954) latitudinal "diasystem" in its broad outline). Some dialects on comparison with this prototypical standard language will be better matches than others. These better matches will form the core dialects and the less good examples will constitute peripheral dialects (for an example of an application of the notion of "nucleus" to dialectology, see Thomas 1990).

The same concept applies to longitudinal variation. Given some nucleus of cohesive grammatical behavior which serves as the prototype for a historical period, further examples of the language can be judged as either central or peripheral examples of this schematic grammar. To the extent that a sample of linguistic material is a better or worse example of one such nucleus, to that extent the language of the sample can be treated as belonging to the linguistic period defined by that nucleus.

The question arises how does one determine what set of cohesive behaviors will serve as a nucleus for a period or language? In the case of dialects, the standard language serves this function. In the case of historical variation, we will actually have several nuclei, one for each period we wish to establish. The starting point will always be the language contemporaneous with the analyst, since that will always be the only linguistic behavior that he has the ability to test by experiment and speaker feedback, whereas knowledge of previous stages cannot be so supplemented by speaker intuition. Taking, then, the current language as the initial nucleus, the analyst can compare textual data to isolate any other nuclei of relatively cohesive behavior. The analyst will encounter increasingly peripheral examples of the nucleus as he moves back into time and eventually can discover that the examples of the language are 'gravitating' towards a new nucleus as prototype. In the same way, a dialectologist can discover that a group of dialects forming a gradient start gravitating towards a different standard as prototype, so that he can see that a dialect is a relatively better example of Dutch than of German, i.e. is closer to the nucleus of standard Dutch than of standard German.

Thus the examples of the language in question form a gradient that shades into adjacent nuclei serving as prototypes. In the case of longitudinal variation, there is a definite directionality to the peripheries (since time is a bi-directional dimension). There will of course be examples of language behavior which are relatively central to the nucleus; these will be judged fairly contemporaneous in nature. But the peripheral examples of a particular nucleus will be treated differently, depending on which nucleus they gravitate towards. If the example gravitates towards a prior nucleus, then the behavior is judged **archaic;** if towards a later one, then it is **innovative.** Thus we can have examples of a language which can be categorized as belonging to a certain period (e.g. takes the nucleus of Middle English as a prototype), but is a peripheral member of the category in the direction of a historically prior nucleus (e.g. a text with Old English features is archaic and gravitates towards the Old English prototype, even though it is relatively a better example of Middle, than of Old, English). At some stage we reach a transition zone, like the Rhenish Fan, where innovative old behavior differs only slightly from archaic new behavior. Thus we are able to capture the analysis and intuitions of philologists concerning the often mixed nature of texts and relate it to a general theory of category structure which also applies to non-historical variation in the language.

One might question at this point whether there is any inherent limit to the number of historical periods one can posit for a language, since one can treat any attestation grouping as constituting a nucleus and find others which are better or worse examples of this prototype. The answer would appear to be no, there is no inherent limit, though there are factors which make some corpora better nuclei than others. The problem can again be put into perspective by a comparison with dialects. It is possible to carry respect for variation so far as to treat each ideolect as a different grammar. There is, of course, some justification in this. But it is also clear that ideolects can be grouped into larger categories by their having some (perhaps schematic) body of grammatical behavior which serves them as a prototype. In the same way, one could isolate one period of a language for every year of its being spoken. But a grammar of 1975 English is of no greater use to general theory than is the grammar of John Smith's English. In longitudinal studies, individual annual phases clearly gravitate towards a more stable nucleus of grammar.

In practice, though, we are more likely to isolate some nuclei as distinct periods than others. In comparing different texts, we are more likely to see a change in prototype if the differences between them are more structural, i.e. more central, in nature than transient or peripheral. One approach would be to adopt a rating of various types of linguistic behavior by their degree of grammaticalization similar to what some dialectologists have proposed. Such scales are, however, subject to debate as to their constituents and precise ordering (Chambers & Trudgill 1980:115-16). Nevertheless, as a first approximation I would argue for the scale of:

morphology > syntax > lexicon > phonology.

Phonological differences, as reflected in historical texts, are subject to synchronic dialectal variation, orthographic traditions, or just plain bad spelling; they are not so reliable as signals of structural change. Although phonological changes often fuel other changes, it would be more difficult to isolate a separate linguistic period solely on the basis of apparent phonological variations in texts. Lexicon can vary greatly within what we would consider a single historical period and a word stock is not such a basic, systemic feature. Syntactic differences, however, are more structural in nature and morphology the most grammaticalized of all linguistic behavior. Using such a scale we can determine the relative weight of the differences between phases: if we see a significant number of sufficiently weighty differences, a different nucleus is posited. One need not be detained at this point with the specifics of such a scale. The potential for defining an adequate rating of linguistic behaviors is sufficient to support a further exploration of periodization based on identifying significant nuclei, however we eventually decide the relative weighing of ''significant''.

Just as planets and stars, though invisible, can be detected by their effects on other, visible bodies, so a nucleus can be discerned when we see it acting as a prototype towards which texts gravitate. To the extent to which a significant number of texts gravitate towards a proposed nucleus (i.e. appear to use it as a standard), to that extent we are justified in positing it as a valid period of the language.

We can also consider certain extra-linguistic factors. As with dialects, we ought not banish these non-linguistic considerations entirely from our definition process, but they must be given their proper weight (e.g. the notion of a 'standard' for any language is always based on extra-linguistic criteria). Extra-linguistic boundary markers sometimes make convenient divisions for analysis and one could just as conveniently take a single attestation grouping as the basis of an analysis as one could a set period. Extra-linguistic factors can help us isolate nuclei, but the establishment of periods must be made on a comparison of the *linguistic* features.

Finally, as with dialect studies, we can also rely somewhat on speaker intuitions about periods. Though not a faultless criterion (Chambers & Trudgill 1980:3-4), speakers' claims of intelligibility can help define dialectal nuclei. The same can be true of historical periods. This is a useful heuristic for making preliminary divisions, as long as it is followed up with a comparison of proposed nuclei for degree of differences and directions of gravitation.

4.0 I have said very much in an abstract, theoretical vein. At this point I want to discuss briefly some facts from the history of Welsh which illustrate how these notions might be applied.

Welsh is traditionally divided into Old, Middle and Modern periods; the division between the last two is usually placed in the early 15th century (see Morris Jones 1913: 6-7; Jones 1979:19-24). Although, on an intuitive level, Welsh has undergone less change in the same time than has English, there are a number of important differences between its periods. A question arises on examination of the texts whether it is proper to treat all samples of Welsh from 1400 onward as representing one phase of the language, or in my terms, is there only one grammatical nucleus in this period or more? The answer lies, as I have said, in comparing the prototypes of behavior for this period with the nuclei of present-day and Middle Welsh to determine any differences. All aspects of the grammar would be examined in a full investigation; here I present data relating to a single aspect of syntax: word order.

Middle Welsh texts differ markedly from contemporary Welsh in the predominant use of a word order called the Abnormal Sentence, whereby some constituent besides the verb occurs first in the sentence. See Mac Cana (1973); Fife & King (1991) for a detailed discussion. Contemporary Welsh contrasts in having VSO as the predominant word order pattern. Although verb-initial is the most common word

Table 1. Three Highest Frequency Word Orders in 3 Classical Welsh Texts

CI	E	G
A	A	A
V	V	V
NS	NS	PS

order in present-day Welsh, in Middle Welsh it occurs in less than 5% of main clauses (Poppe 1989:50). Obviously the two constitute different nuclei of behavior as relates to this basic feature of syntax. But under the traditional scheme of periodization, Modern Welsh, the period in which present-day Welsh belongs, stretches back to 1400. This begs the question of whether an examination of the data indicates that the Welsh of say 1600 had the same grammatical prototype as that of 1900. An examination of the data on the topic of word order shows that the nucleus for 1600 was rather different from that of Middle Welsh, but not equivalent to present-day use either.

The data here are based on research reported in Fife (1991). Examining three texts coming from the period 1540-1620, shows that the three word orders of highest frequency are (1) Adjunct-initial; (2) Verb-initial (3) Subject-initial (either nominal or pronominal). See the overall frequencies presented in Table 1.

Comparing this with the top three orders from several Middle Welsh texts indicates a rather different prototype for word order. See Table 2.

As we can see, in Middle Welsh two-thirds of the texts had adjunct-initial as first choice; as second choice, one-third had one order, one-third another and the last third split between two other orders. Most of the texts had subject-initial order as third choice. Thus the later texts were less variable in word order choice and consistently had verb-initial clauses as second choice. None of the Middle Welsh texts had verb-initial as a major word order. The position of verb-initial clauses on the list of word orders in Middle Welsh averaged 5.6 out of 6 possible orders (see Table 3), while in the later texts, verb-initial order was consistently second. The percentage of verb-initial clauses in Middle Welsh ranged from 0% up to 9%; in the later texts it ranged from 19% to 29%. The average frequency in the later texts was nearly six times higher than in Middle Welsh. Furthermore, a special word order called 'inversion' (see Fife 1986) was rather frequent in Middle

Table 2. Three Highest Frequency Word Orders In 6 Middle Welsh Texts

KO	B	M	CL1	BR	BM
I	A	PS	A	A	A
A	NS	I	NS	I	NO
NS	PS	A	PS	NS	PS

Table 3. Frequency of Verb-initial Order in Middle and Classical Welsh Texts, By Position on Frequency List and Percentage

	Middle Welsh						Class. Welsh		
	KO	B	M	CL1	BR	BM	CI	E	G
Frequency of V over 6 Possible Word Orders	6	6	6	6	6	4	2	2	2
Percentage of V	9	4	0	0	2	9	19	26	29
	MW Average: 4%						CW Average: 25%		

Welsh, being second in frequency for one-third of the texts and most frequent in one text (the oldest of the group). There was not a single instance of inversion in the later texts.

This investigation shows that the texts from the period 1540-1620, what I call Classical Welsh, differed significantly from those in the Middle Welsh period. There can be no doubt that the Classical Welsh texts look to a different grammatical prototype from the earlier texts. As regards word order, the prototype for Classical Welsh is more similar to Modern Welsh than is the Middle Welsh nucleus. Present-day Welsh is typified by having more consistent word order than Middle Welsh, much more reliance on verb-initial order and far less use of inversions. All these features are evident in the Classical Welsh texts. However, the Classical Welsh nucleus also differs from that of contemporary Welsh in that adjunct-initial sentences are still more frequent than verb-initial, as in Middle Welsh, and occurring with about the same degree of frequency.

We can detect a phase of linguistic behavior in Welsh where main clause word order was radically different from the previous period, but also not identical to current behavior. On the basis of these data it might be best to posit a new period between Middle and Modern Welsh, one which forms a coherent body of linguistic behavior which acts as a prototype for classifying the samples of language shading back into the Middle Welsh nucleus or forward towards the contemporary grammar. How well we can establish Classical Welsh as an important nucleus of behavior in the history of Welsh depends on how numerous and how significant is the amassed set of differences from previous and subsequent behavior. This task awaits future research.

5.0 Language variation is a linguistic concept which is readily handled in prototype theory, but which creates problems for classical categorization requiring a binary choice as to membership in one category or another. Historical periods, like synchronic varieties and dialects, are best viewed in a fashion which allows for gradient shading from one point of definition to another. A recognition of gradients and nuclei gets us away from the arbitrariness of extra-linguistic boundary markers and allows us to focus on factors which are most directly relevant to linguists, actual language behavior.

James Fife
6655 La Jolla Blvd. #11
San Diego, CA 92037

REFERENCES

Bailey, Charles-James. 1973. *Variation and linguistic theory.* Arlington: Center for Applied Linguistics.

Bloomfield, Leonard. 1933. *Language.* New York: Holt, Rinehart & Winston.

Boas, Franz. 1940. *Race, language and culture.* New York: Macmillan.

Bourcier, Georges. 1981. *An introduction to the history of the English language.* Tr. C. Clark. Cheltenham: Stanley Thornes.

Brachin, Pierre. 1985. *The Dutch language: A survey.* Tr. P. Vincent. Leiden: E.J. Brill.

Chambers, J., and Peter Trudgill. 1980. *Dialectology.* Cambridge: Cambridge University Press.

Dixon, R. 1982. *Where have all the adjectives gone?* Berlin: Walter de Gruyter.

Fife, James. 1986. "The semantics of GWNEUD inversions." *Bulletin of the Board of Celtic Studies* 32:92-126.

———. 1991. "Some constituent-order frequencies in Classical Welsh prose." In Fife, James, and Erich. Poppe (eds.) *Studies in Brythonic word order.* Amsterdam and Philadelphia: John Benjamins. Pp. 251-73.

———, and Gareth King. 1991. "Focus and the Welsh Abnormal Sentence: A cross-linguistic perspective." In Fife, James, and Erich Poppe (eds.) *Studies in Brythonic word order.* Amsterdam and Philadelphia: John Benjamins. Pp. 81-153.

Francis, Winthrop. 1983. *Dialectology.* London: Longman.

Jones, R. 1979. "A brief history of the Welsh language." In Stephens, Meic. (ed.) *The Welsh language today.* Llandysul: Gomer Press. Pp. 18-31.

Lakoff, George. 1987. *Women, fire and dangerous things.* Chicago: Chicago University Press.

Lockwood, William. 1972. *A panorama of Indo-European languages.* London: Hutchinson & Co.

Mac Cana, Proinsias. 1973. "Celtic word order and the Welsh 'abnormal sentence'." *Ériu* 24:90-120.

Morris Jones, John. 1913. *A Welsh grammar.* Oxford: Clarendon Press.

Newman, Stanley. 1967. "Classical Nahuatl." In McQuown, Norman. (ed.) *Handbook of Middle American Indians.* Vol. 5. Austin: University of Texas Press. Pp. 179-99.

Nist, John. 1966. *A structural history of English.* New York: St. Martin's Press.

Petyt, K. 1980. *The study of dialect.* London: Andre Deutsch.

Poppe. Erich. 1989. "Constituent ordering in *Breudwyt Maxen Wledig.*" *Bulletin of the Board of Celtic Studies* 37:43-63

Pyles, Thomas. 1971. *The origins and development of the English language.* 2d ed. New York: Harcourt, Brace and Johanovich.

Thomas, Alan. 1990. "Networks, nuclei and boundaries in areal dialectology." In Ball, Martin. et al. (eds.) *Celtic linguistics.* Amsterdam and Philadelphia: John Benjamins. Pp. 25-39.

Weinreich, Uriel. 1954. "Is a structural dialectology possible?" *Word* 10:388-400.

Whitney, William. 1889. *Sanskrit grammar.* Thirteenth reprint [1973]. Cambridge: Harvard University Press.

Williams, Joseph. 1975. *Origin of the English language.* New York: Free Press.

SETSUKO KIYOMI

Animateness and shape in classifiers*

Abstract. Scholars like Adams and Conklin (1973), Denny (1976), and Allan (1977) have maintained that 'animateness' and 'shape' are fundamental categories for classifiers. The hierarchical relation between these categories, however, has not been fully examined. The aim of this paper is to discuss whether there is an implicational universal between them. First classifiers are grouped into two types (i.e. 'free morpheme classifiers' and 'bound morpheme classifiers'), and then their semantic properties are examined, focusing attention on those classifiers which are semantically differentiated via 'animateness' and 'shape.' Our findings indicate that there is no implicational universal that languages must include these two categories since some languages do not include shape-based classifiers and others do not have any classifiers of animateness. We propose, then, that only one of these semantic properties is an obligatory feature of every classifier system.

1. Introduction. This article deals with the investigation of the semantics of classificatory markers for nouns which include 'numeral classifiers.' More precisely, we will examine not only what Dixon (1986) calls 'noun classifiers' and 'noun classes', but also so-called classificatory verbs and intra-locative classifiers.

The importance of the above-mentioned various types of classificatory markers for nouns or 'classifiers' (which will be used as a general term throughout this paper) lies in the fact that one makes use of them to categorize entities. In other words, 'classifiers' are assumed to represent a classification of the world. If we find out which semantic category is the most basic, we may be able to understand better how one arranges and interacts with the world. It has been claimed that 'animateness' and 'shape' are the most basic semantic categories for classifiers. In the present paper, we claim that there is no hierarchical relationship between these two semantic categories.

2. Definition. In this section, we briefly look at three definitions of 'classifiers': Schachter (1985), Dixon (1986), Allan (1977). We then propose our own definition.

2.1 Schachter's definition. There is a strong tendency to use the term 'classifiers' only to refer to 'numeral classifiers'. For instance, Schachter takes this position, defining classifiers as follows: ". . . words which are required by the syntax of a certain language when a noun is also modified by a numeral" (1985:39). This is to say that he uses the term 'classifiers' to denote only 'numeral classifiers.' His argument concerning 'classifiers' suggests that there is what can be called a 'classifier language' which must utilize classifiers for all the nouns. He implies that English, having expressions like "two cups of tea" or "five ears of corn", cannot be called a 'classifier language', because most of the nouns in English are directly modified by numerals.[2]

2.2 Dixon's definition. Dixon (1986), first of all, claims that we ought to make a clear distinction between what he terms 'noun classes', which include most gender systems, and 'noun classifiers', which include numeral classifiers. He regards these as two prototypes, into which all kinds of classifiers are grouped: if classifiers belong to a grammatical system then they fall under noun classes, while if classifiers are composed of a lexical set, they are noun classifiers. 'Noun classes' consist of obligatory grammatical systems marked as prefixes, as in the Bantu languages, or articles, as in French or German. 'Noun classifiers' are separate lexemes, and in many languages, classifiers are necessary only when a noun is modified by a numeral.

Dixon (1986:106) presents the following three criteria with which we can distinguish the systems of noun classes from sets of noun classifiers:

(1) (a) SIZE
(b) REALIZATION
(c) SCOPE

As for (a) SIZE, 'noun classes' involve small and finite sets, while 'noun classifiers' have a great number of systems. Secondly, 'noun classes' are always realized as a closed grammatical system such as an inflection. They may form a morphological unit with number, case, tense, etc., such as French articles wich are fused morphemes showing both noun class and definiteness. In Dyirbal, one morpheme indicates three things: noun class, case, and visibility (i.e., whether the entity in question is in view or not). On the other hand, 'noun classifiers' are always realized as free forms: they are separate morphemes which are never fused with number, case, tense, etc. Thirdly, Dixon states that the scope for 'noun classes' is beyond the noun word. That is, if an affix on a noun indicates some class for the noun, this affix may also apply to other words, in the NP. For example, in Swahili, an affix on a noun applies to all the words in an NP (i.e. demonstrative, numerals, and adjectives) and also to the verb. By contrast, 'noun classifiers' are never referred to by anything outside the NP where they co-occur with the specific noun. If we include predicate classifiers in 'noun classes',[3] each prototype proposed by Dixon would consist of the languages represented as follows:

(2)

Noun Classes	Noun Classifiers
Bantu languages Latin German French Dyirbal Navaho	Languages with numeral classifiers (ex. Thai, Cambodian, Japanese, etc.) Jacaltec

Craig criticizes Dixon's claim, pointing out that Jacaltec is considered one of the classifiers systems at the fuzzy edges between the prototypical 'noun classes' and prototypical sets of 'numeral classifiers', because Jacaltec classifiers appear independently of quantifiers and that the number of classifiers is much less than that of numeral classifiers. She concludes as follows: "The task of identifying classifier languages is not always a matter of clear-cut decision indeed, but rather a matter of definition" (1986a:5).

2.3 Allan's definition. Lastly, we will consider Allan's definition of classifiers. Allan (1977:285) offers two criteria:

(3) a) They occur as morphemes in surface structures under specifiable conditions.

b) They have meaning, in the sense that a classifier denotes some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer).

Allan also introduces four types of classifiers into which languages with classifiers are grouped. They are:

(4) a) numeral classifier languages
b) concordial classifier languages
c) predicate classifier languages
d) intra-locative classifier languages

Numeral classifier languages not only use classifiers for expressions of quantity but also use them anaphorically or deictically. The concordial classifier languages are characterized by the classifying particles affixed to modifiers, predicates, and pro-forms as well as nouns. The predicate classifier type is what Hoijer (1945) called ''classificatory verbs''. For instance, Navaho verbs of motion or location include a classificatory stem that characterizes twelve classes of object. The fourth type, 'intra-locative classifier', consists of noun classifiers affixed to locative expressions which obligatorily co-occur with nouns. Allan says there are only three languages that belong to this type: Dyirbal, Toba, and Eskimo, all of which are unrelated with one another genetically.

2.4 Our definition. We have seen three different definitions of classifiers. Because Schachter's viewpoint is quite limited, it apparently cannot deal with various kinds of classifiers. We will reconsider, then, Dixon's definition and the one proposed by Allan.

We agree with Craig (1986a) in that it is not possible to divide different kinds of classifiers into two clear-cut prototypes as Dixon maintains. Although Dixon asserts that each of these two types has specific formal characteristics, not all kinds of classifiers fit in with each type completely. He admits that there is no appropriate place for predicate classifiers in his schema. On the other hand, Allan's four types do not offer a place for Jacaltec classifiers which are realized as free morphemes but are used without numerals. Since these two approaches have disadvantages, we must develop a more reasonable schema that is applicable to any kind of classifiers. For our schema, then, it is proposed first that Allan's two criteria presented in (3) above are used for our definition of classifiers. Second, we will divide classifiers into two ''types'' (not into two ''prototypes'') based on one of the criteria given by Dixon: 'realization' or 'grammatical status' will be adopted as our single criterion. In this way, we consider that there are five kinds of classifiers including Jacaltec and predicate classifiers and divide these five kinds of classifiers into two groups depending on their morpheme status: whether they are realized as free morphemes or bound morphemes. Moreover, it is proposed that the term 'classifier' is to be used as a general term to cover two types of classifiers. The diagram shown below represents our division with specific kinds of classifiers that fall into each type:

(6)

CLASSIFIERS

Free Morpheme Classifiers	Bound Morpheme Classifiers
Numeral Classifiers Non-Numeral Classifiers	Concordial Classifiers Predicate Classifiers Intra-locative Classifiers

In the following section, each type of classifier in the above diagram will be examined: the free morpheme classifiers are dealt with first, followed by the investigation of the bound morpheme classifiers. We will see the overall view of the semantic features found in each kind of classifier, with special attention to the two semantic categories, 'animateness' and 'shape'.

3. Free morpheme classifiers. As is explained in the previous section, we have divided the 'free morpheme classifiers' into two types: 'numeral classifiers' and 'non-numeral classifiers'. We deal with numeral classifiers first, and non-numeral classifiers (=Jacaltec and Yidiny) will be treated next.

3.1 Numeral classifiers. Numeral classifiers may be found everywhere in the world's languages; they are especially widely spread in Asian languages. We will discuss the semantics of the Asian numeral classifiers discovered by Adams and Conklin, and draw some concrete examples from Burmese, Japanese and Chrau.

3.1.1 Asian languages. Adams and Conklin (1973) studied numeral classifiers in 37 Asian languages including those found in the Malayo-Polynesian, Austro-Asiatic, Mon-Khmer, Sino-Tibetan, Altaic, Dravidian, and Indo-Aryan families. Their three main findings are as follows:

(7) a) Classifiers based on 'animateness' are always to be expected.

b) 'Shape' functions as the important factor to classify objects.
c) A number of classifiers are based on the idea of 'function'.

The first finding, animate/inanimate distinction is regarded as the most basic categorization in the numeral classifier construction. These classifiers separate humans from animals in the following three ways:

(8) a. human/non-human
b. animate/inanimate
c. human animate/non-human animate/inanimate

Moreover, Adams and Conklin maintain that it is often the case for many languages with the human/non-human distinctive classifiers to have more than one class for human.[4] For example, according to Becker, Burmese has a distinction between human/non-human and there are four classes for humans: the highest human class corresponds to the classifier for Buddhas. The following diagram illustrates the five-level social hierarchy (Becker 1975:116):

(9)

human				non-human
'shu' Buddhas	' 'pa' deities saints monks	' 'u' people of status scholars	'yauʔ' ordinary humans	'kauŋ' animals ghosts dead bodies depraved people children

Japanese, on the other hand, presents an example of the opposite case since it has only one classifier for humans: 'nin' or 'ri' stands for the single classifier for humans,[5] while there are several classifiers for non-humans, some of which are given below:

(10) 'hiki'- - - - - animals, insects, fish
'too' - - - - - large animals (e.g. horses, cows)
'wa' - - - - - birds
'bi' - - - - - fish
'hai' - - - - - squids[6]

As to the second finding, there are three basic shapes which are 'length', 'flatness', and 'roundedness'. In Japanese, for example, 'hon'[7] corresponds to a classifier for 'length', 'mai' expresses 'flatness', and 'ko' refers to 'roundedness'. We should pay attention to the fact that 'animalness' is sometimes overridden by the semantic property of 'shape'. For example, Chrau, a Mon-Khmer language, classifies animals into classes for shapes as well as a class for animals only (Adams and Conklin 1973:4):

(11)	a) a class for larger animals:	cow deer
	b) a class of long thin things:	stick, fish, pole, pencil, knife, lizard, snake
	c) a class of flat things:	mat, cloth, small gong, clothing, turtle
	d) a highly unusual class containing:	fishtrap, deer, horse, dog

3.1.2 Proto-Thai. Let us have a very brief look at what Delancey (1986) has to say about Proto-Thai. The Thai languages form a family that has three branches: South-western (SWT), Central (CT), and Northern (NT). He speculates that the following list shows strong candidates for the earliest Thai classifiers: the first three are attested throughout the family; the last two are found only in CT and SWT (Delancey 1986:447):

(12) *tue^{A7} - - - - - animals
*ʔan^{A} - - - - - inanimates
*sen^{C} - - - - - one-dimensional

---

*oŋ^{A} - - - - - humans
*phïən- - - - - flexible two-dimensional

Delancey states that except for *phïən, we find that the other four classifiers conform to the cross-linguistic tendencies proposed by Adams and Conklin (1973) that the first distinctions made in numeral classifiers are between 'animates' and 'inanimates' and between 'humans' and 'non-humans'. The other distinction is made on the feature of 'shape'. Since '*sen^{C}' which stands for 'one-dimensional' is attested throughout the family, we can consider this as evidence for the primacy of one-dimensionality over two- and three-dimensionality.

Having looked at numeral classifiers, we have acknowledged that two semantic categories play main roles: ‘animateness’ and ‘shape’. It is not possible, however, to determine their hierarchy at this point.

3.2 Non-numeral classifiers.

3.2.1 Jacaltec. Craig (1986b) examines the noun classification system of Jacaltec,[8] a Mayan language spoken in the western highlands of Guatemala. Jacaltec noun classifiers consist of twenty-four classes which are all transparent in meaning. She attributes this complete semantic transparency to the fact that only concrete objects take noun classifiers and that there is no general or residual classifier for objects which do not share semantic properties. Then we should restate that not all nouns in Jacaltec take noun classifiers but all the nouns which can be classified in the language must take a classifier preceding them. According to Craig, unclassified nouns are:

(13) a) abstract words such as events or happenings: games, race, story, letter
b) locative nouns: road, church
c) unidentified materials and mixed substance: coca cola, food, garbage, smoke, cloud, star

Craig (1986b:266–67) bases her semantic analysis of classifiers on the assumption advocated by Denny (1976) and divides the 24 Jacaltec noun classifiers into two subsystems: a) social interaction, and b) physical and functional interaction:

(14)

SUBSYSTEM I:
SOCIAL INTERACTION

1. cuman- - -male deity
2. cumi7- - -female deity
3. ya7- - - - -respected human
4. naj- - - - -male non-kin
5. ix- - - - - -female non-kin
6. najni7an- -young male non-kin

SUBSYSTEM II:
PHYSICAL AND
FUNCTIONAL INTERACTION

13. no7- - - - -animal
14. metx’- - -dog
15. te7- - - - -plant
16. ixim- - - -corn
17. tx’al- - - -thread
18. tx’an- - - -twine

7. ixni7an- -young female non-kin
8. ho7- - - - -male kin
9. xo7- - - - -young male kin
10. ho7ni7an- -female kin
11. xo7ni7an- -young female kin
12. unin- - - - -infant
19. k'ap- - - - -cloth
20. tx'otx'- - -soil/dirt
21. ch'en- - - -rock
22. atz'am- - -salt
23. ha7- - - - -water
24. k'a7- - - - - -fire

The semantics of the social interaction classifiers require social rank parameters such as age, respect, divinity, and the inherent semantic notions of sex and kinship. On the other hand, the second subsystem of noun classifiers organizes the world of animals, plants, and concrete objects which humans interact with. For the concrete objects, the classification is based on the semantic feature of material; namely, it is based on the substance out of which the objects are made. Thus Craig maintains: ". . . assignment to a class is done on the basis of the knowledge by the Jacaltecs of the constituting substance of the objects" (1986b:273).

We have seen Jacaltec classifiers and become aware of their remarkable difference from numeral classifiers. One big difference is that in Jacaltec, there are no classifiers based on 'shape' at all. In spite of the fact that Jacaltec classifiers do not require numerals to co-occur with nouns, they can be considered as counterexamples to Adams and Conklin's suggestion that 'shape' is the primary semantic category for classifiers. Another fact worth noticing is that Jacaltec has as many as twelve classes for humans. If we compare them with the four Burmese classes discussed in 3.1.1, it is obvious that Jacaltec has a much more elaborated system for referring to people.

3.2.2 Yidiny. We now move on to Yidiny, an aboriginal language spoken in Australia. It is Dyirbal's northern neighbor, but as Dixon (1982) mentions, it is quite different from Dyirbal in phonology, morphology, and syntax. The classifiers in Yidiny are optionally used, and they do not require a co-occurring numeral or demonstrative.

Dixon divides Yidiny classifiers into two categories on the basis of semantics. The category called 'inherent nature' includes fourteen classifiers and the other called 'function/use' consists of six as shown below (Dixon 1982:191–93):

(15)

Classifiers based on the INHERENT NATURE of their referents	Classifiers based on the FUNCTION or USE of their referents
'waguja- - - -man	'minya'- - -edible animal
'bunya'- - - - -woman	'nayi'- - - - -edible plant
'bama'- - - - -person	'buna'- - - - -drinkable liquid
'maggum'- - -frog	'bilmba'- - -habitable place
'munyimunyi'- -ant	'gugu'- - - - -purposeful noise
'jugi'- - - - - -tree	
'baji'- - - - - -canoe	
'narra'- - - - -wine	
'buri'- - - - - -fire	
'jabu'- -solid inanimate matter other than stone or wood	
'gala'- - - - - -spear	
'bundu'- - - -bag	
'baji'- - - - -canoe	

We should note that Yidiny, like Jacaltec, does not possess any classifiers based on 'shape' either. The classifiers of 'animateness' are found as is expected.

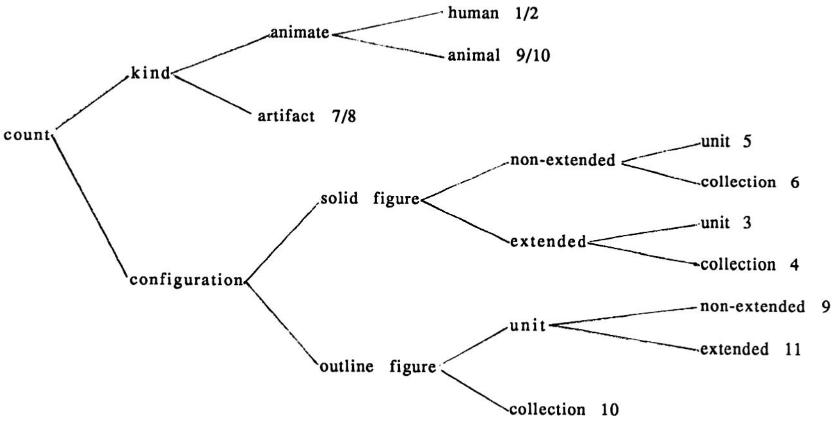
Having made a brief survey of Asian numeral classifiers and of the two non-numeral classifiers, we have observed that 'animateness' is used in all of the free morpheme classifiers, but 'shape' is not. It is therefore reasonable to assume at this stage that 'animateness' be placed highest in the hierarchy of semantic categories as far as free morpheme classifiers are concerned.

4. Bound morpheme classifiers. There are three kinds of bound morpheme classifiers: 1) concordial, 2) predicate, and 3) intra-locative.

4.1 Concordial Classifiers—Proto-Bantu. We will take up Proto-Bantu classifiers as representing concordial classifiers. Denny and Creider (1986) argue that the Proto-Bantu noun prefixes which are regarded as classifiers constituted a semantic system. That is, their claim is that each prefix was associated with a particular characteristic meaning. They have discovered that many noun prefixes expressed configurational or shape meanings.

They first separate count nouns from mass nouns; we will discuss only the former here. The next diagram illustrates how they divide classifiers (i.e. prefixes) according to the semantic feature(s) of the items that the count nouns refer to (Denny and Creider 1986:219):

16)



As is shown in the above, the first division is made by the two semantic features: 'kind' and 'configuration'. With regards to 'kind', two major divisions are made between 'animate' and 'artifact'. Humans belong to class 1/2 and most animals belong to class 9/10. Those animals which are not classified in class 9/10 are categorized in terms of 'shape' or under 'artifact'. Thus, for example, a 'spider' belongs to class 5 which semantically indicates solid figure, non-extended, and a 'frog' belongs to class 7/8: frogs are regarded as despised objects which can be understood as a natural extension of 'artifact'.

Next we deal with 'configuration'. Configurational classes are further divided according to whether solid shape or outline shape is used for the basis of classification. Within each of these pairs, one finds a contrast between 'extended' and 'non-extended' configurations. The term 'extended' means the relative length in one or two dimensions. On the other hand, 'non-extended' is positively characterized as rounded or humped. The contrast between 'outline figure' and 'solid figure' refers to a contrast between objects which show clear edges so that an outside and an inside are different from each other and objects which do not. Under the extended outline figure class (11), items that are curved such as ‘crust’, ‘fingernail, ‘rib’, etc. would be included. Under the non-extended outline shape class (9), items that are rigid and flexible containers like ‘pot’, ‘drum’, ‘house’, ‘skin garment’, etc. would be included. We have ‘body’, ‘lip’, ‘river’, ‘arrow’, ‘leg’, etc. as representing extended solid objects (3 & 4). Non-extended solid figure (5 & 6) includes objects such as ‘egg’, ‘stone’, ‘sun’, ‘nose’, ‘cheek’, etc. We have made a quick survey of the Proto-Bantu concordial classifier system observing that there is a human/non-human distinction and that several classifiers are related with configurational meanings. Thus both ‘animate’ and ‘shape’ classifiers are found in Proto-Bantu, and it is suggested that ‘shape’ may have been primary.

4.2 Predicate classifiers.

The predicate classifiers seem to be limited to the Athapaskan languages. The Navaho classifiers are handled first, then two other Athapaskan languages, Chipewyan and Hupa will be discussed.

4.2.1 Navaho.

In the Athapaskan languages, the predicate classifiers are characterized as having the structure of prefixes plus a classificatory verb stem which varies depending on which noun class the theme of the verb belongs to. That is, if the verb is intransitive, the verb stem varies in accordance with the noun class of the subject. If the verb is transitive, the object of the verb determines the verb stem. Hoijer (1945) discovered twelve such verb stem variants in Navaho. They are normally conjugated with the ‘si-’ or ‘ni-’ perfective marker. Only the forms with ‘si-’ are given below: all of the various verb stems shown mean ‘lie’ (Hoijer 1945:15):

(17)	a.	Round object	sìˀą́
	b.	Long object	sìtą́
	c.	Living being	sìtį́
	d.	Set of objects	sìníl
	e.	Rigid container with contents	sìką́
	f.	Fabric-like object	síłcò·z
	g.	Bulky object	sìžó·d
	h.	Set of parallel objects	sìžò·ž
	i.	A mass	sìʒà·ʔ
	j.	Wool-like mass	sìʒ·l

k.	Rope-like object	sìlá
l.	Mud-like object	sìχé·ˀ

We can see that the above Navaho classifiers include a single class for 'animateness' (i.e. living things) and three 'shape' classifiers (i.e. round object, long object, and rope-like object). The difference between what Hoijer calls 'long object' and 'rope-like object' is that the former signifies rigidness and the latter, flexibility.

4.2.2 Chipewyan predicate classifiers. Another Athapaskan language, Chipewyan, was investigated by Carter (1976) who used elicitation frames with bilingual informants. Carter could elicit as many as ten different patterns of verb stems which include three classifiers for animate objects:

(18)	a.	da - - - 'awake beings'	a frog, a spider, a sitting bear, a sitting person, a beaver, etc.
	b.	tį - - - 'sleeping beings'[9]	a person, a sleeping baby, a sleeping bear, a girl, etc.
	c.	łti - - - 'dead beings'	a dead person, a bear carcass, a dead dog, a raw fish, etc.
	d.	ˀą- - - 'inanimate solid objects'	a lake, an ax, one dollar, a coil of rope, a knife, a stone, an orange, etc.
	e.	kę- - - 'liquids' e.g.	some mud in a container, mud, blood, boiling water, coffee, milk
	f.	dzáy - - 'granular masses'	a pile of sand, a pile of sugar, instant milk, fish, eggs, etc.
	g.	la - - - 'rope-like object, objects in sets, or plurality of objects'	a rope, a book, several arrows, playing card, eyeglasses, veins, etc.
	h.	tą- - - 'stick-like objects or empty containers'	an airplane, a bow, an empty box, a canoe, a chair, a firewood, etc.

i.	łtą - - - ‘containers with contents’	a box with stuff in it, one can of beer, a pack of cigarettes, etc.
j.	ltšuθ) - - - ‘fabric-like objects’	a calendar, parka pants, a sheet of paper, a book, one glove, etc.

The classification of ‘animateness’ is quite interesting in spite of some unclarity (e.g. there is a question why ‘awake beings’ must be sitting instead of standing). Animate beings are divided by the criterion of being ‘dead’ or ‘alive’ first, and then alive beings are distinguished between ‘sleeping beings’ or ‘awake beings’. Additionally, it is necessary to consider whether there are shape-based classifiers in Chipewyan or not. We should recall that Asian languages have classifiers of three basic shapes: ‘length’, ‘flatness’, and ‘roundedness’.[10] It is possible to regard the Chipewyan “-tą̓” as corresponding to length’, “-ItšuΘ”, to ‘flatness’, and “ʔą̓” to ‘roundedness’. Consequently, the Chipewyan predicate classifiers and Asian numeral classifiers have much in common in terms of semantic categories, although the subdivisions of ‘animateness’ are rather different.

4.2.3 Hupa. Now let us examine what Haas (1967) discusses in her comparative study of Yurok numeral classifiers and Hupa predicate classifiers.[11] Haas states that the comparison of these two kinds of classifiers reveals strikingly similar semantic bases in both kinds of classifiers. She concludes that, in spite of the differences in syntactic manifestation between Yurok (= numeral classifiers) and Hupa (= predicate classifiers), they represent similar semantic categories. The diagram below shows what they have in common (Haas 1967: 360):

(19)

Yurok Numeral Classifiers	Hupa Predicate Classifiers
Human beings Animals	Living beings (humans and animals)
Straight or stick-like objects	Long or stick-like objects
Round objects	Round objects
Rope-like objects	Rope-like object (or several objects)

We have dealt with three Athapaskan languages with the purpose of finding out what semantic properties are involved in the predicate classifiers. It has been observed that both 'animateness' and 'shape' are among the semantic categories identified in the predicate classifiers.

4.3 Intra-locative classifiers. As mentioned in 2.3, there are only three languages identified as having intra-locative classifiers: Dyirbal, Toba, and Eskimo.

4.3.1 Dyirbal intra-locative classifiers. The classifiers of Dyirbal, an aboriginal language of Australia are realized in what Dixon (1972) calls 'a noun marker', which involves three kinds of information; location, case, and noun class. In this language, a noun must always be preceded by an appropriate noun marker. We will consider only the nominative case.[12] The classifier is represented as the suffix of the noun marker:

(20) Class I: -l
Class II: -n
Class II: -m
Class IV: -∅

Three different forms appear in the first element of the noun marker, which indicates visibility and proximity:

(21) 'bala-'- - - -visible and there
'yala-'- - - -visible and here
'*ŋ*ala-'- - - -not visible but audible

Therefore we have the following four nominative forms of the noun class referring to 'visible and there':

(22) Class I: bayi[13]
Class II: balan
Class III: balam
Class IV: bala

These four kinds of intra-locative classifiers of Dyirbal were scrutinized by Dixon (1982) and Lakoff (1986). Dixon describes a brief version of the Dyirbal classification of objects as follows (1982:178):

(23)	(I)	*bayi*	men, kangaroos, possums, bats, most snakes, most fishes, some birds, most insects, the moon, storms, rainbows, boomerangs, etc.
	(II)	*balan*	women, bandicoots, dogs, playpuses, echidna, some snakes, some fishes, scorpions, anything connected with water or fire, sun, etc.
	(III)	*balam*	all edible fruit and vegetables and the plants that bear them, honey
	(IV)	*bala*	parts of the body, meat, bees, wine, yamsticks, some spears, most trees and vines, grass, mud, stones, noises and language, etc.

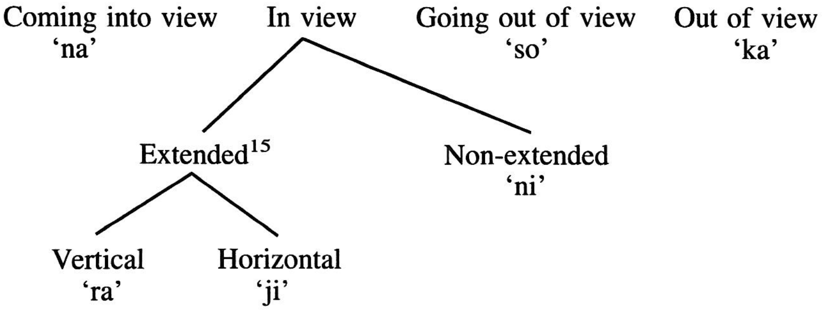
At a glance, the Dyirbal classifiers seem to be gender markers, because 'men' and 'women' are represented by the respective classifier. Besides, we have seen that in Jacaltec non-numeral classifiers, human males and females have independent classifiers. There is a great difference between the Dyirbal classifiers and the Jacaltec ones, however: in Dyirbal not only animals but also natural things such as the moon, rainbows, fire, the sun, etc. take the same classifiers that are used for 'men' and 'women'. Although we can agree with Lakoff (1986) in that the first and the second classes might represent human males and females respectively at a very abstract level, neither of these classes is actually limited to humans nor animates. It is proper, then, to claim that Dyirbal does not have classifiers based on 'animateness'. The third class limits its co-occurring nouns to 'edible things'. The fourth class includes 'everything else', so that it is impossible to find the common semantic property in it: at least we can say that the fourth class includes 'inanimate objects'. Therefore, except for the third class *balam,* which has a particular semantic property, none of the other three categories show a specific semantic property. Since there is no 'animate' classifier nor 'shape' classifiers, it is evident that they are definitely different from the other classifiers we have handled so far. Thus, it is possible to regard them as consisting of a gender system by reason of the opaque semantic correspondence between the classifier and its accompanying nouns. Surveying 'Toba' and 'Eskimo' intra-locative classifiers next, we will be able to know whether all the intra-locative classifiers are regarded as consisting of a gender system or not.

4.3.2 Toba and Eskimo. Klein (1979:85) mentions that Toba has six classifier morphemes occurring as noun prefixes: only the feminine forms are preceded by the feminine marker 'ha-'. Klein analyzes the meaning of each morpheme as follows:

(24) 'na' - - - - - coming into view
'ni' - - - - - in view, non-extended[14]
'ra' - - - - - in view, vertical
'ji' - - - - - in view, horizontal
'so' - - - - - going out of view
'ka' - - - - - out of view

Unlike the Dyirbal classifiers in which the locative markers and the classifiers have independent forms, the Toba classifiers and locative markers cannot be morphologically distinguished from each other. We may assume that the Toba classifiers are morphologically merged with locative markers. Denny claims, therefore, that four locations are distinguished in the Toba locative system and only the entities *in view* are classified, so that there are only three classifiers. His diagram below shows that there is no morpheme to designate visibility itself; the three classifiers denote that the entities referred to are 'in view' (Denny 1979b:102):

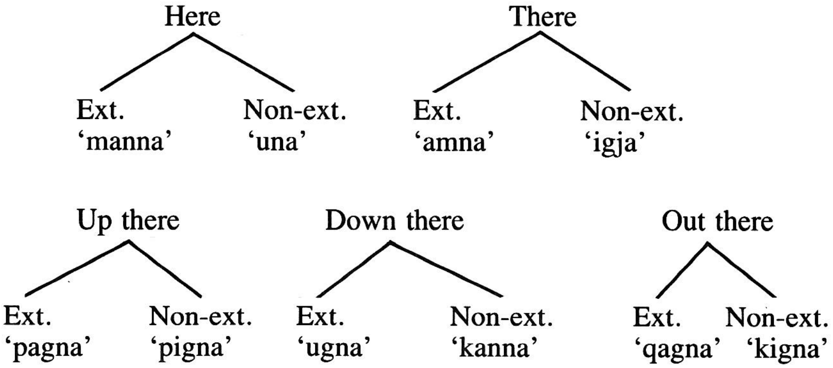
(25)



For instance, ‘‘jilames’’ (=‘table’) has a prefix ‘ji’, which indicates that a table is seen and is extended and horizontal. In sum, Toba has three classifiers of ‘shape’; non-extended, vertical, and horizontal.

Denny concisely accounts for Eskimo classifiers. There are six locations: 1) here 2) there 3) up there 4) down there 5) out there 6) on there. Except for entities ‘in there’,[16] the distinction between ‘extendedness’ versus ‘non-extendedness’ is made in five locations. These five pairs of classifiers, which occur as demonstrative pronouns[17], are illustrated below (Denny 1979b:103):

(26)



The above diagram indicates that Eskimo, like Toba, has classifiers morphologically fused with locative markers.

We have argued that Dyirbal classifiers can be regarded as a gender system because of the opaque semantic properties of the classifiers. Toba and Eskimo classifiers, however, cannot be taken for genders. This is because both Toba and Eskimo have classifiers based on ‘shape’: Eskimo classifiers show a single contrast between ‘extendedness’ and ‘non-extendedness’, while Toba has a three-way distinction: ‘vertical’, ‘horizontal’, and ‘non-extended’. In addition, what is more important to note is that no classifiers based on ‘animateness’ are found in these languages. Consequently we can summarize what we have observed regarding the intra-locative classifiers as follows: 1) the Dyirbal classifiers ought to be regarded as a gender system 2) the Toba and Eskimo classifiers, on the other hand, share a semantic property of ‘shape’, but they have no classifiers based on the semantic property of ‘animateness’.

We have seen three kinds of bound morpheme classifiers: 1) concordial classifiers, 2) predicate classifiers, 3) intra-locative classifiers. It has been discovered that both of the concordial and the predicate types include the classifiers of 'shape' and 'animateness'. The intra-locative type, however, possesses only the classifiers of 'shape', and no classifiers of 'animateness' are discerned. Therefore, within the bound morpheme classifiers, 'shape' is considered to be placed higher than 'animateness' in the implicational hierarchy of the semantic categories.

5. Conclusion. The following chart illustrates our findings:

(27)

CLASSIFIERS

Free Morpheme Classifiers	Bound Morpheme Classifiers
Numeral Classifiers (Animate, Shape) Non-Numeral Classifiers (Animate only)	Concordial Classifiers (Animate, Shape) Predicate Classifiers (Animate, Shape) Intra-locative Classifiers (Shape only)

As far as the free morpheme classifiers are concerned, it has turned out that the numeral classifiers include classifiers which are semantically based on 'animateness' and 'shape'. The non-numeral classifiers (i.e. Jacaltec, and Yidny classifiers), however, do not include shape-oriented classifiers. Therefore it was surmised that 'animateness' might be placed higher than 'shape' in the semantic hierarchy of classifiers.

On the other hand, among the bound morpheme classifiers, the intralocative classifiers do not include animateness-based classifiers. Although the classifiers based on 'shape' and 'animateness' are found in the concordial and predicate classifiers, the intra-locative classifiers are only shape-oriented. This might lead us to postulate that 'shape' is in a higher position than 'animateness' in the semantic hierarchy, which contradicts our former claim that 'animateness' is higher.

From the discrepancy between these two assumptions, therefore, we are able to conclude that there is no implicational universal between the semantic categories of 'animateness' and 'shape'. In other words, each of these two semantic categories is regarded as an independent, semantic prerequisite for all kinds of classifiers. Thus, on the assumption that 'animateness' and 'shape' are the two most fundamental and independent semantic categories for classifiers, we can maintain that either the semantic category based on 'shape' or that based on 'animateness' has to be included in every kind of classifier system.

Department of Linguistics
Indiana University
Memorial Hall 303
Bloomington IN 47405

ENDNOTES

*I am grateful to Linda Schwartz, Steven Franks, and Stuart Davis for their valuable comments on an earlier version of this paper.

[1] Dixon's distinction between 'noun classifiers' and 'noun classes' will be accounted for in 2.2.

[2] Schachter does not offer much evidence for his statement that English is not a classifier language. Allan (1977) gives a much more definite explanation in this respect. According to Allan, there are seven semantic categories: (1) material (2) shape (3) consistency (4) size (5) location (6) arrangement (7) quanta. The first five categories are based on the inherent characteristics of the objects, whereas the last two are not. It is important to note Allan's observation that the first five are only found in classifier languages but the last two can occur in non-classifier languages such as English.

[3] Dixon (1986) admits that predicate classifiers (= 'classificatory verbs') found in Athapaskan languages do not fall under either of the two schemata. He suggests, however, they may belong to 'noun classes', for they satisfy some criteria (i.e. they are regarded as a small closed system of stem types).

[4] In addition, they have found that 1) humans are classed according to either social rank or kinship but not both of them, 2) humans are never categorized on the basis of sex alone.

[5] Japanese has two phonological forms for the human classifiers. 'Ri' is used only for 'one' and 'two', while 'nin' co-occurs with 'three' or more.

[6] 'Hiki' is a general classifier for various kinds of fishes including squids. 'Bi' and 'hai' are special forms used for fish and squid respectively: they are used much less often than 'hiki'.

[7] The 'h' in 'hon' phonologically alternates with 'b' and a geminate 'p'.

[8] Jacaltec has three types of obligatory classifiers: the first one is a set of number classifiers represented as suffixes forming a three way opposition; human/ animal/ inaminate. The second type is plural classifiers; it has only two markers. The third type is what Craig (1986b) calls 'noun classifiers' (N CL) which function as free morphemes that immediately precede the noun and can co-occur with other determiners. It is the third type (= N CL) we are concerned with here.

[9] Carter (1976) does not give any explanation as to why 'a person' is included in 'sleeping beings.'

[10] This is discussed in 3.1.1.

[11] Yurok is an Algonquian language and Hupa is an Athapaskan language.

[12] The other case markers are ergative, dative, and simple genitive.

[13] Dixon (1972) points out that the nominative class I is irregular.

[14] By 'non-extendedness', Klein means "no parts sticking out into space" (Klein 1979:94).

[15] Denny (1979b:78) states: "The extendedness variable in noun classifier semantics is the one whose values have most often been glossed with terms like 'round', 'long', and 'flat'."

[16] Denny (1979b) gives two reasons for why the things 'in-there' are not classified into 'extended' and 'non-extended': 1) since 'in there' designates a remote location, the space referred to would be constricted 2) the extendedness variable in Eskimo is much longer than in other cultures.

[17] This does not mean that the Eskimo classifiers are free morphemes. The Eskimo classifiers are merged with the locative markers and the merged forms stand for demonstrative pronouns. Since the Eskimo classifiers cannot appear independently of the locative markers, they are regarded as bound morpheme classifiers.

REFERENCES

Adams, Karen L. and Nancy Conklin. 1973. "Toward a theory of natural classification." *Papers from the 9th Regional Meeting of the Chicago Linguistic Society,* 1–10.

Adams, Karen. 1986. "Numeral classifiers in Austroasiatic." In Craig (ed.), 241–62.

Allan, Keith. 1977. "Classifiers." *Language,* 53:284–310.

Basso, Keith H. 1968. "The Western Apache classificatory verb system: a formal analysis." *Southwestern Journal of Anthropology,* 24:252–66.

Becker, Alton L. 1975. "A Linguistic image of nature: the Burmese numerative classifier system." *Linguistics,* 165:109–21.

Carter, Robin M. 1976. "Chipewyan classificatory verbs." *International Journal of American Linguistics,* 42.1:24–30.

Comrie, Bernard. (1981) 1989. *Language universals and linguistic typology.* 2nd ed. Chicago: The University of Chicago Press.

Craig, Colette G. 1986a. "Introduction." In Craig (ed.), 1–10.

———. 1986b. "Jacaltec noun classifiers: a study in language and culture." In Craig (ed.), 263–93.

———(ed.). 1986c. *Noun classes and categorization: proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983* (Typological Studies in Language, 7). Amsterdam: Benjamins.

Delancey, Scott. 1986. "Toward a history of Tai classifier systems." In Craig (ed.), 437–52.

Denny, Peter J. 1976. "What are noun classifiers good for?" *Papers from the Twelfth Meeting of the Chicago Linguistic Society,* 122–32.

———. 1979a. "Semantic analysis of selected Japanese numeral classifiers for units." *Linguistics,* 17:317–35.

———. 1979b. "The 'extendedness' variable in classifier semantics: universal features and cultural variation." In Mathiot (ed.), 97–119.

———, and Chet A Creider. 1986. "The semantics of noun classes in Proto-Bantu." In Craig (ed.), 217–39.

Dixon, Robert M.W. 1972. *The Dyirbal language of North Queensland.* Cambridge: Cambridge University Press.

———. 1982. *Where have all the adjectives gone?* The Hague: Mouton.

———. 1986. "Noun classes and noun classification in typological perspective." In Craig (ed.), 105–12.

Friedrich, Paul. 1970. "Shape in grammar." *Language,* 46:370–407.

Haas, Mary R. 1967. ''Language and taxonomy in northwestern California.'' *American Anthropologist*, 69:358–62.

Hoijer, Harry. 1945. ''Classificatory verb stems in the Apachean languages.'' *International Journal of American Linguistics*, 11:13–23.

Klein, Harriet E.M. 1979. ''Noun classifiers in Toba.'' In Mathiot (ed.), 85–95.

Lakoff, George. 1986. ''Classifiers as a reflection of mind.'' In Craig (ed.), 13–51.

Mathiot, Madeleine (ed.). 1979. *Ethnology: Boas, Sapir and Whorf revisited*. The Hague: Mouton.

Schachter, Paul. 1985. ''Parts-of-speech systems.'' In Shopen (ed.), 3–61.

Shopen, Timothy (ed.). 1985. *Language typology and syntactic description, vol. 1: clause structure*. Cambridge: Cambridge University Press.

WINFRED P. LEHMANN

Indo-European linguistics today

Abstract. Examination from a typological viewpoint of the linguistic data available in the early IE dialects has provided possibilities for explaining problems left unsolved by our predecessors. Much attention has been devoted to the phonological system, where the positing of laryngeals has led to clarification of the vocalic system, including ablaut, and to reassessment of the obstruent system. But the most significant advances have been achieved by viewing the entire linguistic system in keeping with the understanding of language provided by 'contentive' typology. Reconstructing an early stage of Proto-Indo-European as an active language has, for example, clarified the long-standing difficulty of the relationship of the middle and the perfect. It permits us to account for the various expressions for 'have' in early dialects. It lets us account for the twofold lexical expressions for 'water', 'lie', and so on.

Moreover, examining languages, including proto-languages, for their overall system, and for their sub-systems, provides more secure evidence for aligning linguistic findings with those of archeologists and scientists interpreting their findings than has the piecemeal attention of earlier Indo-Europeanists. For example, scrutiny of the numeral system rules out use of tokens by early speakers of Proto-Indo-European, and accordingly a 'homeland' in the Middle East, where tokens were in use from about 8500 B.C. More important than such individual conclusions is the impetus provided by our deeper understanding of the proto-language to compare our findings with those of other scientists of prehistoric peoples, as in correlating strata in the developing linguistic system with datable artifacts of the third and earlier millennia.

1. Approaches in Indo-European studies. When we examine IE linguistics today we cannot forget that we are working in a long tradition. Moreover, it is important to keep in mind the theories, methodologies and aims of past and current Indo-Europeanists; the importance of one's methodology has once again been brought to our attention by the recent publications in the *Scientific American* and the *Atlantic* that touch on our field while largely concerned with Nostratic and other proposed protolanguages. Our predecessors have provided us with massive handbooks on all facets of language, the lexicon, phonology, morphology, syntax, semantics and pragmatics, or as many handbooks call it, stylistics. (For a comprehensive set of references see Szemerényi 1989.)

The advanced state of our field has determined the strategy of my presentation here. Treatment of new contributions requires such massive attention to facts that a survey might get bogged down in details. In the brief space available here, I first sketch the three successive approaches of the past two centuries; then I discuss some of the unsolved problems. Finally I note the relevance of solutions to some of these problems for current contributions from other fields dealing with IE studies, notably archaeology.

We date the start of IE linguistics with Franz Bopp's monograph of 1816, but his work and that of his successors were influenced by two earlier publications. One was Friedrich Schlegel's treatise 'on the language and the wisdom of the Indians' of 1808; in this treatise Schlegel declared his concern for ''innerste Sprachform.'' Strongly influential, this treatise directed Indo-Europeanists from the beginning to determine what many of us as students referred to as 'inner form' and now may be more frequently referred to as 'underlying form or underlying structure.'

Schlegel's treatise was followed in four years by the massively influential publications of the anatomist, Georges Cuvier. According to his theoretical approach, ''every organized being consists of a unique and integrated system in which the parts are closely interrelated.'' Influenced by Cuvier's insistence on the essence of any convention in its organization, Indo-Europeanists looked for structure in language. We see the results in Bopp's monograph, in Grimm's rules, and so on (cf. Lehmann ed. 1967). We also see Cuvier's theoretical approach maintained in the course of IE studies, probably expressed today most often in the words of Meillet that language is a structure ''où tout se tient''—where everything is interrelated. I stress this attention to theory from the beginning of IE studies because so-called theorists appear from time to time and suggest, if they don't state openly, that Indo-Europeanists and previous linguists do not base their work on theory. The earliest theoretical revolution that I know of during the two centuries of IE studies is Grotefend's, who in 1827 presented to the world a new theory of syntax. I omit reference to more recent productions.

Indo-Europeanists have set out to achieve the aims expressed by Schlegel and Cuvier under three approaches, each of which is prominent roughly in successive periods, though these do not exclude the other approaches. The third, which will occupy most of our attention, builds on the two earlier approaches. I illustrate these briefly through well-known contributions.

2. The systematic approach. The first I date from 1816 to approximately 1900. This approach may be characterized by Brugmann's term 'systematic'. In the preface to his important *Grundriss* Brugmann stated that the time is not yet right to abandon a systematic approach in favor of a historical approach (1897:ix). That is to say, his work is essentially descriptive. To illustrate his attention to system, I may mention that he posits for PIE the set *s sh z zh.* The system of stops at the time consisted of four such members at each place of articulation. In keeping with his theoretical approach he proposed the same set for the fricatives. Although we no longer do, I do not want to leave the impression that Brugmann's theoretical approach detracts from his work. The *Grundriss,* like many other publications of the time and earlier, has not been matched for comprehensive and accurate presentation of the facts at hand then. I might also mention that the approach has been followed by many Indo-Europeanists since; Szemerényi's handbook of 1989 lists eight cases for the IE noun, three to six tenses for the verb, four moods, and so on. By the approach, one determines the structure in a language and provides a systematic presentation.

3. The historical approach. The second approach I label historical; I date it from the sixth decade of the nineteenth century to the fifth of ours. Under this approach linguists set out to relate successive stages of language to one another, and also to reconstruct earlier stages. We may contrast Jacob Grimm's confining his treatment of the consonants to rules that indicate relationships between elements in one language and another, for example, between Germanic voiceless fricatives and Latin voiceless stops, and so on. As Grimm acknowledged, ablaut was too complex for him to treat in this way. Ablaut remained largely obscure until the time of Hirt, who contributed greatly to its clarification by positing successive stages in the vocalic system. A characteristic product of the historical approach is Hirt's *Indogermanische Grammatik* (1921–37). However one regards its details, it represents results achieved by application of the two central methods in historical linguistics that were perfected during this period, the comparative method and the method of internal reconstruction.

But the conclusions achieved by Hirt were superseded through the discovery of new data. When Hittite data were available, it became clear that the IE phonological system as posited by Hirt, and for that matter the morphological and syntactic system as well, were based on inadequate information. Yet the new information was processed initially by the historical approach. As recent monographs and collections of papers like that edited by Bammesberger (1988) indicate, a great deal of energy is still applied in accordance with this approach.

4. The linguistic approach based on typological findings. An achievement often announced for historical study is its provision of explanation, the highest of Whitney's three goals for linguistics—accurately recording the data, describing them, explaining them. While contributing to our understanding of specific problems, for example, the contrast in final consonants of *was* : *were*, the historical approach has its obvious limits. For example, it does not permit us to account for the presence of two words for 'fire' in PIE—the etymon of our word and cognates like Hittite *pahhur,* as well as that of Latin *ignis* and Sanskrit *agni-*. Nor does it lead to an explanation for the absence of a word for 'have' in PIE; there is no straightforward etymology for our own word 'have'. We cannot then reconstruct a PIE etymon from it, nor from Latin *habēre,* nor from Greek *ékhein,* etc., not to mention the totally different equivalents, like the Latin *mihi* est construction. To provide an explanation for such unsolved problems we rely on the third approach, which I call linguistic drawing on typological investigations.

Typological studies made up the most prominent theoretical linguistic activity in the eighteenth century. They have their culmination in Adelung's *Mithridates.* 'In this massive work, Adelung and Vater present central characteristics of numerous languages. Though still of great interest, *Mithridates* had the misfortune of preceding Cuvier's insight that linguistic characteristics must be regarded as elements in a system in which the constituents are closely related to one another. It took almost a century and a half to apply that insight to typological study outside morphology, and to some extent phonology. Trubetzkoy provided major advances in phonology during the second to fourth decades of our century, leading to his major contribution of 1939. Two decades later Greenberg published a seminal essay for syntax (1963). In the meantime Soviet linguists were applying Cuvier's insight to language as a whole, as presented in Klimov's important theoretical works (1977, 1983). For IE studies the central work under this approach is *Indo-European and the Indo-Europeans* by Gamkrelidze and Ivanov (1984).

Contentive typology, in Klimov's designation, deals with content as well as grammar, with meaning as well as form. The two volumes by Gamkrelidze and Ivanov demonstrate the wider attention by allotting more space to the lexicon than to the grammar. Viewing language as a whole, contentive typology identifies two principal types of language. One in which government is central is manifested in accusative and ergative languages. The other, in which agreement rather than government is central, is manifested in active-stative and class languages (cf. Lehmann 1990).

A major advance has been made in IE studies through the recognition that an earlier stage of PIE was an active-stative language; cf. Gamkrelidze and Ivanov 1984, Chapter 5, pp. 267–317, esp. 5.6, pp. 308–11. This advance is the most important recent contribution to IE linguistics; its implications have by no means been exhausted. I list briefly some of the clarifications resulting from that identification, after noting pertinent features of active-stative languages, generally referred to, here also, as active languages.

In active languages, first identified by Sapir (1917), sentences are constructed by pairing active or animate nouns with active verbs, inactive or inanimate nouns with stative verbs. That is to say, there are two large classes of both nouns and verbs, active/animate and inactive/stative. Stative verbs, briefly, are intransitive or adjectival. Active languages then have no adjectives. Moreover, the chief inflection is in the verbal system, and there for active verbs. Since transitivity is not central, active languages do not include a verb for 'have', using instead a locational phrase like the Latin *mihi est* 'to me is' construction. Among examples of active languages are the Dakota and Tunica of North America, the Tupí-Guaraní of South America.

In view of uninformed remarks about language typology, I mention before we precede that I do not posit, nor do I identify, ideal active, or ideal accusative, or ideal ergative languages. Languages change, as has been known for some time. While, in accordance with Cuvier's dicta, the changes are governed by their structure, providing the regulated process labeled drift by Sapir, change may also be brought about by neighboring languages of different type. And as Klimov has proposed, there may be universal drift of class from active to ergative to accusative languages, leaving all languages with residues. Even though languages are not ideal representatives of a specific type, unless one is an inveterate stamp collector in every pursuit, including linguistics, viewing language as an assortment of unrelated facts seems imprudent.

5. Solutions resulting from the assumption that Pre-Indo-European was an active-stative language. When we regard an earlier stage of PIE as active, we solve a number of long-standing problems. One in the noun system I have already mentioned: the presence of two words for 'fire' and also for 'water'; it is paralleled in the verb system by similar pairs, as for 'lie' and 'stand'. While active languages classify words like those for 'person' and 'cow' as animate in contrast with words like those for 'stone' or 'stick', they may have two terms for objects or processes that can be regarded as either animate or inanimate, such as 'fire' and 'water'. We may recall that *ignis* is a masculine noun and that *Agni-* is god of fire; Greek *pũr,* Gothic *fon* on the other hand are neuter. Meillet dealt with the problem resulting from presence of two terms for items like 'fire' and 'water'. He ascribed the use of such animate terms to a semi-civilized mentality (1937:340). However high our regard for Meillet, we always suspect such nonlinguistic explanations for elements of language; our science has developed by progressively eliminating them. Identification of Pre-IE as an active language provides a linguistic explanation. The twofold inclusion of such terms is a residue from the earlier active type. Upon shift to accusative structure, only one term was necessary in a dialect. The term that was maintained was selected in each through a course of events we cannot reconstruct.

Something more fundamental can also be accounted for in the nominal system. The threefold IE gender system is late. Meillet in the *Introduction* posited two genders, animate and inanimate (1937:339). That system continues the earlier active structure.

Meillet then goes on to discuss the cases. On the basis of their forms and uses, he distinguishes between the nominative and accusative on the one hand, and the five other cases on the other (1937: 341–42). But he does not pursue the basis for the different position of the nominative and accusative. We may recall that Uhlenbeck did. In a two-page article of 1901 he proposed that an early system of PIE had two cases, which he labeled Aktivus and Passivus. Their functions he characterized as Agens and Patiens, pointing out that Dakota as well as Eskimo and Basque had such systems. In other words, by comparing the proposed case system of Pre-IE with that of Dakota, Uhlenbeck recognized that both were active languages. It is a simple matter to substitute our terms animate and inanimate for Aktivus and Passivus; the underlying case terms that he used are well-known from recent publications. (For an extensive discussion, see Gamkrelidze and Ivanov 1984:Chapter 5.2)

Uhlenbeck's proposal, while routinely referred to in handbooks, was not incorporated into **their analyses.** Nor did proponents of underlying case forms like Fillmore come to terms with it (1968:13–14).

It is important for us to understand the reason for both the lack of response to his proposal and that of Meillet concerning the early gender system. Neither proposal was linked to the structure of the language. In neither proposal was a language type indicated that might have made the analysis credible. In commenting on the absence of credibility, I do not mean to downgrade either Uhlenbeck or Meillet. What is important is the stage of linguistic study. Noone had proposed an active type in 1901, let alone the importance of dealing with languages in accordance with typological findings. The IE handbooks of the day, and many still, take the structure of IE languages to be somehow right, or to use a favored term in our field 'natural.' Forces like transitivity are supposedly fundamental in language. The welcome attention to active languages has broadened our scope, and our realization of the importance of typological study. It is for this reason that I characterize the third and current period of IE studies as based on a linguistic approach drawing on typological investigations.

Surveying features in the IE nominal system that have remained puzzling we point out briefly the solutions provided by the assumption of earlier active structure.

Adjective are virtually indistinguishable from nouns in PIE; pointing to the difference only in the neuter nominative-accusative, Meillet proceeds to treat them together (1937:252ff.).

Further, we cannot reconstruct a genitive for the proto-language; the forms differ too widely. As explanation, active languages have no genitive; not including the feature of transitivity, they express the relationship through some kind of locative. Genitives were created in individual dialect groups of the IE family as the languages came to be accusative in structure.

In his Proto-Germanic grammar of 1939 Prokosch tentatively explained morphological problems of the first and second persons plural of personal pronouns by comparing the presence of inclusive and exclusive pronouns in some languages. This seemed to him the best way of accounting for the use of the root *we-* in the first person, as in our word *we,* and also in the second, as in Latin *vōs* 'you'. The proposed explanation seemed to be another of the random comparisons found in linguistic publications that can often be exploded by data from other languages. When, however, we observe that active languages distinguish inclusive and exclusive pronouns, Prokosch's observation carries conviction. To use another term favored in our field today, the explanation is motivated. (See now Gamkrelidze and Ivanov 1984, Chapter 5.4.1, pp. 290–92.)

Turning to the verb system, we note that the assumption of earlier active structure solves one of the persistent problems: the similarities between the perfect and the middle, and, further, the additional similarity of the Hittite *hi*-verb system. All three are adaptations of the earlier stative inflection in contrast with the *mi*-inflection that was continued from the active inflection. One of the remarkable characteristics of the IE perfect is the poverty of forms; there are distinctive forms only for the singular. In active languages the stative inflection is also poorly equipped. This formal feature, which applies also to the middle and to the Hittite *hi*-inflection, would be adequate by itself to support the assumption of development from the Pre-Indo-European stative. But the use of the forms also supports the hypothesis. The perfect expresses state, as does the middle frequently. Many of the verbs that are inflected in the *hi*-inflection in Hittite have the same meaning as verbs assigned to the stative system in active languages. In short, our major problem in the verb system is now explained.

We may add briefly that the inability to reconstruct a form for 'have' in PIE is also explained. Lacking the feature of transitivity, active languages do not include a verb for 'have'. In short, major problems in the verb system of the proto-language are solved by assuming earlier active structure much as are problems in the nominal system.

6. Among tasks to be undertaken. I may appear to be suggesting that all our problems in IE studies are now solved. After all, much of what I have discussed is in the Gamkrelidze and Ivanov handbook. By contrast, I would like to assert that the field for many fruitful investigations has just been opened up. When a shift takes place, there are always slips. As we have seen, the stative verb system split up into the perfect system of Sanskrit, that we know as the perfect tense in Latin; in early Greek it became the perfect system, that developed into the perfect tense and then merged with the aorist, etc., while in other dialects we find different alignments. Similarly, the middle soon develops into the passive in many dialects. We now need investigations of the shift from active to accusative structure in each of the dialects. I cite only a few examples.

Our Greek grammars tell us that the verb *hēmai* 'sit' is inflected only in the present system. Further, that 'in place of *hēmai* we find usually *káth-ēmai* in Attic prose and comedy. *káthēmai* sometimes is perfect in meaning (*I have sat, I have been seated*) (Smyth and Messing 1956:216). We now know their explanation; active languages often contain both an active and an inactive item for 'sit', the active meaning 'sit down'; *hēmai* with its restriction to the present system and its occasional perfect meaning continues the Pre-Indo-European stative verb.

In our authoritative grammars we learn as well that *keīmai* 'lie' is regularly used in the present and imperfect instead of the perfect and pluperfect passive of *títhēmi* 'place'. Our current historical perspective, expanded by results of typological study, permits us to understand the interrelationships of the two roots and their inflections. Clarification of such problems and others in the grammar that are more extensive can occupy the energies of Indo-Europeanists for some time.

7. Implications for the areal situation and the homeland. We may turn to other matters with implications beyond linguistics. Contentive typology requires us to examine the lexical structure as well as the grammatical structure of language. Examining one such structure, the cardinal numerals, we find that the early numerals are based on demonstratives and the higher are obscure in etymology. Our word *one,* Gothic *ains,* IE **oy-no-* is related to the demonstrative in Latin *is, ea, id* 'this one close to me' (Pokorny 1959:286). Our word *two,* Greek *dúo,* IE **dwōw* I have related to a demonstrative particle meaning 'further', as in Hittite *tuwa* 'distant'. And *three,* Latin *trēs* I have related to the root **ter-* as in Sanskrit *tiráḥ* 'beyond' (1990). The other numerals are difficult to explain, except that *five* is related to Hittite *pankus* meaning 'whole'. I may still mention that Hittite attests for 'four' the term *mi-ya-u-e-es,* which Neu relates to the root **mey-,* with the meaning 'lessen' (1987:176–77). If *miyawes* is the inherited word, it implies the lesser hand, that is, four fingers. We may conclude that the early IE speakers counted by pointing to objects, and used a term for the whole hand when they had exhausted their fingers. The words for 'six', 'seven', 'eight', and 'nine' I assume were borrowed.

When we recall the cultural situation of the fourth millennium and earlier, we observe that the process of enumeration was tied to the use of tokens. These stood for items of economic importance such as fleeces, jars of oil, and other items that were used in trade and ultimately for taxation. Schmandt-Besserat has documented their development from approximately 8500 B.C., and their replacement about 3200 with precursors of cuneiform symbols (1992). That is to say, she has documented the steps towards development of writing as well as the procedures for administering the increasingly complex economic affairs of early civilization.

We cannot identify many of the languages of this pre-literate time. But we know where tokens were used. They have been excavated throughout the Middle East, including south-central Turkey, the home of the Indo-Europeans as assumed by Renfrew. When we hypothesize the numeral system of the languages of that area, we may assume that it contained numeral classifiers, not simply demonstratives for pointing to objects. For obviously the use of tokens required counting the number of sheep, of fleeces, of jars of oil, and so on. The number systems by this hypothesis would be quite unlike that of PIE. By a further assumption, Renfrew's proposed homeland is unlikely.

We may review briefly another point. In some respects Pre-Indo-European lacks characteristics of an active language so that the structure may be the result of influence from another language. Curiously, Northwest Semitic also has characteristics of an active language. Its perfective is also defective, while the imperfective is taken to be the older paradigm. It contains few or no adjectives. Noun inflection is poor, as is gender marking. The common characteristics between IE and Northwest Semitic that some linguists have pointed to may be the result of borrowing from a third language. We look forward to elucidation from archeological discoveries, first of texts but also of other data.

Anthony and Brown have now published their findings on the location of the earliest discovered bridling of horses. The findings are based on almost unbelievably painstaking studies, as on the wear of horse's teeth caused by rope bridles, the examination of quantities of teeth, and finally the localization of the place where the first worn teeth are found. The place is Dereivka in southern Russia; the period approximately 4000. It is about at this period that the terminology for wagonry, and for the horse, was introduced into PIE. As Specht summarized it fifty years ago, that terminology is inflected thematically (1944). Accordingly the wagonry terminology is younger than words for nature, for the family, for social arrangements, for simple living. The findings support the assumption that the homeland of the Indo-Europeans, at least in the fifth millennium, is in southern Russia.

Linguistic findings in this way correlate with archeological findings, though we look forward to far more data. With luck we will find more Hurrian material, so that we will have fuller knowledge of a language that influenced Hittite and possibly also PIE. We are learning more about the dissemination of IE speakers eastward, as into India. And we are more assuredly reconstructing a proto-language that is not a composite of Sanskrit and Greek. If we deal with that language realistically, we can make progress in identifying relationships, both genetic and areal, with other languages of the fifth millennium and earlier. IE linguistics today then provides many opportunities for research, and for further knowledge of the development of the language family whose members are most widely spoken today from origins whose disclosure will illuminate the early stages of human civilization.

Linguistics Research Center
The University of Texas at Austin
PO Box 7247
Austin TX 78713-7247

REFERENCES

Adelung, Johann Christoph. 1806–16. *Mithridates, oder allgemeine Sprachenkunde*. I-III. Completed by J. S. Vater. Berlin: Voss.

Anthony, David W. and Dorcas R. Brown. 1991. "The origins of horseback riding," *Antiquity* 65:22–38.

Bammesberger, Alfred. ed. 1988. *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Winter.

Bopp, Franz. 1816 *Über das Conjugationssystem der Sanskrit-Sprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprachen*. Ed. K. J. Windischmann. Frankfurt a.M.: Andreä.

Brugmann, Karl. 1897–1916. *Vergleichende Laut- und Stammbildungs- und Flexionslehre der indogermanischen Sprachen*. 2d ed. = *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. Strassburg: Trübner.

Cuvier, Georges. 1812. *Recherches sur les ossemens fossiles de quadrupèdes*. Paris.

Fillmore, Charles H. 1968. "The case for case," *Universals in linguistic theory*. Eds. Emmon Bach and Robert T. Harms. New York: Holt, Rinehart & Winston. Pp. 1–88.

Gamkrelidze, Thomas V. and Vyacheslav V. Ivanov. 1984. *Indoevropejskij jazyk i indoevropejcy*. Tbilisi: Publishing House of the Tbilisi State University.

Greenberg, Joseph. 1963, 1966. "Some universals of grammar with reference to the order of meaningful elements," *Universals of language*. Ed. Joseph Greenberg. Cambridge: MIT. Pp 73–113.

Grotefend, August. 1827. *Grundzüge einer neuen Satztheorie* . . . . Hannover: Hahn.

Hirt, Hermann. 1921–37. *Indogermanische Grammatik*. 7 vols. Heidelberg: Winter.

Klimov, Georgij A. 1977. *Tipologija jazykov aktivnogo stroja*. Moscow: Nauka.

———. 1983. *Principy kontensivnoj tipologii*. Moscow: Nauka.

Lehmann, Winfred P. 1990. "The current thrust of Indo-European studies," *General linguistics* 20:1–52.

———, ed. 1967. *A reader in nineteenth-century historical Indo-European linguistics*. Bloomington: Indiana University Press.

Meillet, Antoine. 1937. *Introduction à l'étude comparative des langues indo-europénnes*. Paris: Hachette.

Neu, Erich. 1987. "Zum Wortschatz des Hethitischen aus synchroner und diachroner Sicht," *Studien zum indogermanischen Wortschatz*. Ed. Wolfgang Meid. Innsbruck: Institut für Sprachwissenschart. Pp. 167–88.

Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch.* Bern: Francke.

Prokosch, Eduard. 1939. *A comparative Germanic grammar.* Philadelphia: University of Pennsylvania Press.

Renfrew, Colin. 1987. *Archaeology and language: The puzzle of Indo-European origins.* London: Cape.

Ross, Philip E. 1991. "Hard words," *Scientific American* 264(4):138–47.

Sapir, Edward. 1917. Review of C. C. Uhlenbeck, *Het passieve karakter van het Verbum Transitivum of van het Verbum Actionis in Talen van Noord-Amerika. International journal of American linguistics* 1:82–86.

Schlegel, Friedrich. 1818. *Über die Sprache und Weisheit der Indier.* Heidelberg: Mohr and Zimmer.

Schmandt-Besserat, Denise. 1992. *Before writing.* Austin: University of Texas Press.

Smyth, Herbert Weir and Gordon M. Messing. 1956. *Greek grammar.* Cambridge: Harvard University Press.

Specht, Franz. 1944, 1947. *Der Ursprung der indogermanischen Deklination.* Göttingen: Vandenhoeck & Ruprecht.

Szemerényi, Oswald. 1989, 3d ed. *Einführung in die vergleichende Sprachwissenschaft.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Trubetzkoy, Nikolay S. 1939. *Grundzüge der Phonologie.* Prague: Travaux du Cercle Linguistique de Prague 7.

Uhlenbeck, C. C. 1901. "Agens und Patiens im Kasussystem der Indogermanischen Sprachen," *IF* 12:170–71.

Wright, Robert. 1991. "Quest for the mother tongue," *The Atlantic* 267(4):39–68.

GARLAND CANNON

Sir William Jones, language families, and Indo-European

Abstract. This paper will reevaluate Jones's famous 1786 formulation and his other findings that essentially founded modern linguistics. Building on ideas about language relationships that had been voiced for two millennia, but casting these and his own interdisciplinary findings into the form of a theory, his 1786 formulation was a refined, tested form of an experimental hypothesis based on a systematic phonological and morphological comparison of words and inflections primarily from Sanskrit, Latin, Greek, and various stages of Persian, anticipatory of Grimm. He rejected divine explanation and chance and explained massive language similarities on the basis of an earlier source in common. Heavily influenced by Linnaean taxonomy, he postulated that languages can change and die, like organisms, and innovationally grouped similar languages into separate families, rather than to leave them unexplained as thousands of unrelated languages around the world. The family to which he gave his greatest attention was Indo-European, projecting branches like Celtic, Indic, Germanic, Iranian, Hellenic, Italic, etc. He mapped out the IE family, people, homeland, and even migration patterns to help explain language change. His theory is one of the most important in the history of ideas and ultimately marked the end of the conception that God gave Hebrew to Adam as a direct gift of the first language, initiating the separation of religion from linguistics, which could then move toward science.

Despite the towering achievements of 19th-century linguists like Rasmus Rask and Jacob Grimm, and of 20th-century scholars like Ferdinand de Saussure and Leonard Bloomfield, the most famous and accomplished linguist probably deservedly remains Sir William Jones (1746–94). As justification for such renown, scholars and lay persons usually cite his "sprung from some common source" formulation in 1786, the influence of his literary translations and Oriental literary imitations upon Western Romanticism, his introduction of the world of Sanskrit culture to a startled West, and individual discoveries like his identifying the Greek Sandrocottus as the Maurya emperor Chandragupta I, which enabled Western historians to work back and forth from this known Greek date so as to integrate Indian history into a world perspective.

In the history of ideas, Jones's 'common source' formulation is by far his most important contribution to world knowledge. It is worthwhile to reconsider that formulation primarily in terms of its also presenting the concept of language families, in a programmatic statement of what later became known as the Indo-European family of languages and its numerous progeny over millennia, a source for which he borrowed the Sanskrit word *aryas* 'noble' and provided the modern English spelling *Aryan*. His important elaborations of this concept are scattered among passages in various essays written in India before his early death in 1794. As he viewed language research almost exclusively as a means of assisting human beings, rather than viewing linguistic revelations particularly as valuable in their own right, he never composed an integrated study of language families and their migration routes and dispersion. Thus, one must study all of Jones's Indian writings in relation to his common-source passage so as to construct his overall view, particularly in view of the multidisciplinary sources that contributed to Jones's primarily linguistic formulation. Such a study shows that he was the father of Indo-European studies, the development of which, concomitant with the needed comparative method, gave rise to modern linguistics.

Five conclusions emerge from a study of Jones's total Indian writings, the last four of which required a fine linguist to integrate the nonlinguistic materials into an overall generalization. First, a scholar was needed who had a breadth of learning and scholarship particularly in history, anthropology and archaeology (such as they were at that time), and literature so as to mesh together focal areas and migration routes along with a reasonable chronology of historical events and nations. Second, a polyglot was needed to recall or find similarities among the needed numbers of languages. Third, data from exotic languages were needed, so as to force this polyglot into new ways of thinking about phonological and morphological structures. Fourth, in view of the massive 18th-century ethnocentricity and racism, where already through Gibbon, Adam Smith, and other scholars there was a view of Western civilization as deriving mainly from Roman and Greek cultures as perhaps the maximum human accomplishments, these languages had to be genealogically related to the extensively studied European languages and thus almost certainly had to be Asian. Despite Jones's linguistic genius, had he migrated to America as he contemplated doing because of his opposition to the British military and political oppression of the American colonies, even his most brilliant work with any American Indian language probably would have had little or no impact on the rise of modern linguistics, just as any work with a European, but non-Indo-European language like Hungarian or Etruscan would have had similarly limited influence. Fortunately for the world, Jones became a judge in the Bengal Supreme Court in 1783 and was able to study Sanskrit intensively for nine years. It was primarily a by-product of this work in Sanskrit, including his translation of the 12,608 words in his still-unpublished and unedited manuscript of the *Amarakośa* (in the Texas A&M University Library) that led to his projecting the concept of an Indo-European family among other language families.

The fifth conclusion is that he needed Sanskrit data in developing his concept, partly because 19th-century Continental linguists could not make major advances until they had mastered the Pāṇinean approach and had absorbed a knowledge of Sanskrit phonetics into their work (Emeneau 1955). By 1779 Jones was already one of the most famous translators and language scholars in the world, having built his reputation primarily upon Arabic and Persian. His *Grammar of the Persian language* (1771) was already in the third of its ultimately nine English editions and two French ones, and had already influenced Oriental grammars by George Hadley, John Richardson, and Nathaniel Brassey Halhed. It was on the way toward becoming one of the most famous, influential grammars of an exotic language ever to be done.

A fine scholar from the outset, Jones was aware of the numerous pronouncements about language relationships over the centuries. One possible source that he apparently did not know was Dante, of whose work he mentions only the *Divina Commedia*. Nonetheless, Dante's *De vulgari eloquentia* began with an elaborate discussion of God's giving Adam the sacred language of Hebrew in the Garden of Eden, a language that was thought to persist until the building of the Tower of Babel, whereupon each of the various artisan types had a different language—the architects, stone movers, stone cutters, etc. All these vernaculars were thus derived from Hebrew, which continued to be spoken by the people who did not participate in building the tower and so were not scattered throughout the world. Dante's explanation used time as a sufficient cause of language change, which explained the later divergence of Latin into Spanish, French, and Italian, as one of three groups of related languages in Europe. Italian, praised as the most illustrious of all, further divided into at least fourteen vernaculars. Dante recognized that a major dialect may be expressed by several microdialects, each in one city, as in his example of Bologna. Then the great European scholar Joseph Justus Scaliger (1540–1609) projected a four-fold division of some European languages on the basis of words that mean 'god'—*deus* (Latin, etc.), *gott* (English, etc.), *theos* (Greek), and *bog* (Russian) (see Mallory 1989:9–10). The problem of mutally unintelligible languages was biblically resolved, as people spoke in tongues to express the wonderful works of God.

Jones may have been a subscriber to James Parsons's tedious, 419-page *Remains of Japhet: Being historical enquiries into the affinity and origin of the European languages* (London, 1767). Parsons's total reliance on biblical genealogy to attempt to relate some languages obscured some quite perceptive observations. In mainly arguing the thesis that 18th-century Irish was directly descended from Japhetan, Parsons ably used 1,000 cognates to prove the affinity between Irish and Welsh. He attempted to tie fourteen European languages together (Greek, Latin, Spanish, Italian, French, Irish, Welsh, German, English, Dutch, Danish, Swedish, Polish, and Russian), and even added a few cognates from Persian and Bengali. But even a formulation like "I count the *Irish* and *Welsh* to be sister-dialects of the *Pelasgian* language" (p. xii) depended upon a Genesis-type basis, and serious errors like the inclusion of Hungarian as an Indo-European langauge further obscured Parsons's linguistic contributions.

Jones may have known the fragmentary generalizations about linguistic genealogical relationships made by Filippo Sassetti, Père Gaston Cœurdoux (who used the phrase "une commune origine" in relation to grammatical and phonological relationships among selected Sanskrit, Latin, and Greek words), and Père du Pons, though he never mentioned them in his extensive correspondence. As a classical scholar, Jones knew the almost two-millennia-old observations about commonalities between Greek and Latin, though the obviously massive Latin borrowing from Greek considerably beclouded the real relationship between Latin and Greek. Having studied Spanish, Portuguese, and French, he observed without particular reaction many commonalities among these three languages, just as his work with German constantly turned up other cognates. Yet in his seven books concerned with languages in the period 1770–74, there is no really meaningful observation about such commonalities.

It took a letter from the Polish prince and scholar Adam Czartoryski in November 1778 to catalyze Jones into formulating his views about language relationships. Yet Jones's answer in February 1779 to Czartoryski's pointed questions was undistinguished in its perceptions and implications. He said that he could not explain with certainty "how so many European words crept into the Persian language," but cited statements in Procopius's *History of the wars* (A.D. 550) about "the great intercourse, both in war and peace, between the Persians and the nations in the north of Europe and Asia, whom the ancients knew by the general name of Scythians." Jones then referred generally to the conviction of "many learned investigators of antiquity . . . that a very old and almost primæval language was in use among these northern nations, from which not only the Celtic dialects, but even the Greek and Latin, are derived." He illustrated with two sets of cognates from Greek, Latin, and Persian, but concluded that "these researches are very obscure and uncertain" and far less agreeable and valuable than the study of a Háfiz ode or an elegy by Imru'u 'l-Qays (Jones, *Letters,* 1:285). So besides discouraging the wealthy Polish general from a promising genealogical study, Jones still had no insightful ideas on the subject.

Now let us transport ourselves to Bengal in September 1785, when Jones engaged a brilliant pundit at the ancient Sanskrit university at Nabadwip and began to immerse himself in Sanskrit religious manuscripts. The reason for his quickly almost fanatical devotion to Sanskrit was typical of him: although he was a powerful member of a colonial administration in a conquered Asiatic land, he strongly opposed European injustices against Indians. European administrators, particularly the judges, could not govern Indians (whether Hindu or Muslim) fairly, according to their own laws and customs, unless there were English versions of those laws. The court pundits and maulvis were often ignorant of the ancient *smritis* and were sometimes bribable. Though Charles Wilkins was being permitted to study Sanskrit in Banaras at full salary by the East India Company, he was already manifesting a literary and linguistic bent that precluded his becoming a Sanskrit clearinghouse for Jones's many legal and religious scholarly questions. Wilkins had no interest in the *Mānava-Dharmaśāstra,* which evidently contained many of the answers. So Jones turned intensively to Sanskrit, in the process being plunged into Pāṇinean grammar with its explicitly stated segmental morphemes that helped him to intuit cognates, primarily from Persian, Latin, and Greek, along the same analytical lines and thus helped suggest his "common source" formulation. Ironically, the posthumous publication of Jones's English translation under the title of *Institutes of Hindu law:* or, *The ordinances of Manu* (1794) had no great impact on justice in British India or on overall posterity, whereas the importance of his conception of the

Indo-European family can hardly be exaggerated. Equally ironically, he was not particularly impressed by his development of this concept, as it held no direct improvement for his fellow beings.

The welter of 18th-century prescientific thinking, perhaps characterized somewhat earlier by a brilliant Newton trying to date the creation of Adam and Jason's search for the Golden Fleece, is a separate, huge topic. This welter and particularly the strong biblical underpinning vastly complicated Jones's perception of an intriguing, emerging generalization: there were many resemblances among Sanskrit, Persian, and familiar European languages. Excepting mainly Herder, scholars were still hoping to locate the Garden of Eden. It was still widely believed that God gave Hebrew to the first human being as a mature and direct gift, and that Hebrew was thus the source of all languages. Jones was a devout but independent Christian, and would have liked to trace the contemporary inhabitants of the Middle East back to a single source, but this desire for commonality would never have biased him into positing a linguistic explanation before weighing the evidence. The reconstruction of his procedure is similarly a lengthy subject (see Cannon 1990:241–46 and 1991). We will simply summarize that he reached his integrative conception independently in a scientific method anticipatory of comparative linguistics. He packed his explanation into a lucid, dramatic few lines in his long Third Anniversary Discourse in 1786 that would revolutionize Western thinking and ultimately that of the world:

> The *Sanscrit* language, whatever be its antiquity, is of a wonderful structure; more perfect than the *Greek*, more copious than the *Latin*, and more exquisitely refined than either, yet bearing to both of them a stronger affinity both in the roots of verbs and in the forms of grammar, than could possibly have been produced by accident; so strong indeed, that no philologer could examine them all three, without believing them to have sprung from some common source, which, perhaps, no longer exists: there is a similar reason, though not quite so forcible, for supposing that both the *Gothick* and the *Celtick*, though blended with a very different idiom, had the same origin with the *Sanscrit;* and the old *Persian* might be added to the same family, if this were the place for discussing any question concerning the antiquities of *Persia* (*Works* 3:34–35).

Clearly, this formulation indicates that Jones had found the analytical method which separated actual cognates of free forms and their inflections from similarities caused by accident or borrowing, a method that Parsons had earlier hinted at. Either of these two causes would have required a separate explanation for each set of similar forms, whereas one genealogical explanation accommodated all such similarities exclusive of accidental resemblances. Jones and his colleagues in the innovational Asiatic Society that he had founded in Calcutta were botanizing widely over the subcontinent, and his intensive study of Linnaeus had led him to apply to languages the then still-visionary goal of developing a single comprehensive classification of *flora* that would permit all botanists to agree on the identity and description of any specimen worldwide within a binomial nomenclature. Buffon, Linnaeus, and some early geologists were observing that at least some flora and some fauna were not immutable. Jones completed the picture by applying this idea to languages, in a mutability distinct from biblical causes. Languages could die or change just as an organism could die or change, and that linguistic process could be as revealing as Buffon's and Linnaeus' biotic studies were. Perhaps linguistics needed a scientific model to move into modern modes of thought.

Two points are critical about Jones's formulation. First, he succintly laid out the skeleton of the Indo-European family, which would later enormously excite Rask, Grimm, Bopp, Verner, etc., who would then develop modern linguistics in the process of subsuming phonological changes within a guiding principle. They worked with the Germanic languages, and it was vital that Jones's skeleton would permit the use of rich Germanic data. Had he gone to a British-controlled area of Arabia, say, and had made a similar programmatic conception of Semitic, there might not have been linguistic fame for a Grimm or a Friedrich von Schlegel. The second point is that Jones explained language similarities in terms of a *family,* in perhaps the first use of this old word to describe a genealogical kinship of languages, though echoing Parsons's earlier *sister-dialects*. Without this concept, his explanation could have been a random one to describe a chance grouping, so that one might then project thousands of isolated pairs or small groups of genetically related languages, or perhaps just one chance grouping and thousands of unrelated languages around the world. Languages, like organisms, had genetic relationships among themselves. The concept went far beyond the groupings of Dante, Scaliger, Parsons, and Cœurdoux. Jones stripped away all religious encumbrances and convincingly used only one example from each of several Indo-European branches. What he did was to elevate his explanation into a theory that could generally explain all such linguistic similarities worldwide, even if the details continued to elude scholars, as in the modern problems of Australian aborigines' languages and those of some American Indians. Then he exemplified this theory in scattered passages in succeeding essays, by extending it to non-Indo-European Asiatic languages and projecting various other language families.

These essays provided 19th-century Germanic scholars with the final clues for the scientific study of languages on the basis of historical, comparative change. Jones divided a language's lexicon into two kinds of items. One kind consisted of a dependable set of basic data containing the names of parts of the body, common actions, material elements, etc. The second kind included names of exotic things and actions, which were much more likely to be borrowed than were everyday items and so should not be used for linguistic comparisons. Jones insisted that scholars should use etymological and inflectional data under rigid conditions, rejecting the kind of folk etymology often used by his contemporaries Lord Monboddo, Jacob Bryant, and John Horne Tooke, though he did not stress inflectional data to the degree that modern scholarship would require.

Mainly, Jones gave future linguists the vital concept of language families and historical change. He refined and expanded this concept through the use of terms like *common stock* and *common parent*. His letters contain numerous references to Sanskrit as the *sister* of Latin and Greek, just as he placed Ethiopic in the Semitic family as the *sister* of Arabic and Hebrew. His rejection of the biblical explanation is exemplified by his stress on the differences between Hebrew and the Indo-European languages, which were so great that he said that they could never have been part of the same family. Indeed, he gave equal prominence to Arabic and Hebrew. He articulated the subfamilies derived from Indo-European, though of course not mentioning Hittite and Tocharian, which were discovered later. Thus, Hellenic gave rise to Aeolic, Doric, Attic, and perhaps Phrygian. Indic provided Sanskrit, the Prakrits, Bengali, Hindustani, and Romany, though Romany's differing vocabulary troubled Jones. Other branches were Slavic, Iranian, Romance, and Germanic, with Germanic diverging into mutually unintelligible subbranches containing English, German, Swedish, and Gothic, respectively. He postulated a family containing Turkish, Mongolian, and other tongues, all of which could be traced back to a single earlier source that presumably no longer existed. The Finno-Ugric family contained at least Lappish, Finnish, and Hungarian.

Jones said comparatively little about the Indo-European language, though he argued that it was *not* a primitive, inadequate system of communication spoken by Asian savages. He speculated that the Indo-Europeans cultivated no liberal arts, and they were certainly nonliterate. But they "formed a variety of dialects, as their tribes were variously ramified" into the Indo-European subfamilies, whereupon Indo-European passed out of existence (*Works* 3:202). Even more strikingly, their Sanskrit descendants lived in a Golden Age that may have surpassed the glory that was Greece and the grandeur that was Rome: the people who spoke Sanskrit were "splendid in arts and arms, happy in government, wise in legislation, and eminent in various knowledge" (*Works* 3:32). It was this relating to the greatness of Sanskrit culture that galvanized Europe.

Such praises plus Jones's translations of fine Sanskrit literature like the *Śakuntalā* completely contradicted Europeans' anthropological, philosophical view of Indians as a ferocious, savage people who committed suttee and infanticide and made sacrifices to Kālī, savages who logically should not have had a great past. Europeans had known of the Sanskrit language since Purchas's reference in his *Pilgrimage* (1613), but knew nothing about the riches of Sanskrit culture. To Europeans, the Indians could not be logically connected to the Europeans, who perceived themselves as having evolved out of original savagery into civilization. Now Jones had proved a direct linguistic (and distantly ultimately ethnological) relationship with Europe for the descendants of the Prakrit speakers. These "savages" had composed the *Mahābhārata* and the Vedas; and their descendants had made the seemingly unrustable Iron Pillar with its Gupta poetic inscription, had built the brooding Black Pagoda at Kanarak, and had richly cultivated astronomy, philosophy, music, etc. Jones and his Asiatic Society colleagues communicated these discoveries to Europe and America in their organization's *Asiatick Researches,* the volumes of which were widely pirated and translated in Europe and also pirated in America, overshadowing the Royal Society's *Philosophical Transactions* and *Histoire de l'Académie Royale des Inscriptions et Belles-Lettres*. By 1800 Jones's true view of classical India had been accepted by Europe and America, and Sanskrit culture was on its way toward assimilation as an important part of world knowledge. It was this Sanskrit—Indo-European—European connection, once it was perceived and appreciated, that wrought such a radical change in Western thought.

Finally, what did Jones's theory of languages families bequeath to posterity? Clearly, he did not believe that a given parent robot-like emitted descendant languages equally, and certainly not simultaneously. The Indo-Europeans migrated westward, except for the Indians, helping shape the false bias that these were *Indo*-Europeans, though Jones placed their focal area as Iran and never hypothesized that they lived in India. Isolation, time, contact with different biota, and other factors would lead to perhaps mutually unintelligible languages within five hundred years. While he may have yearned to link all language families back to one super-family like Nostratic, his stern warnings about the need for massive data, rigid methodology, and empirical evidence anticipated the problems that have plaqued modern efforts to prove distant relationships among Semitic, Indo-European, Ural-Altaic, Dravidian, and some other language families. Nor did he attempt to place the Indo-Europeans in the Garden of Eden, or to propose Sanskrit as the parent tongue of the Indo-European languages, as Monboddo, Schlegel, and others unwarily did. No doubt Jones's phenomenal scholarly stature and cogency played their part in keeping theologians from attacking him, as they attacked Darwin seventy years later.

As has been said, the idea of language relationships had been bandied about for two millennia. What Jones did was to put such otherwise isolated facts together with many other linguistic and non-linguistic facts into an integrated, ultimately interdisciplinary theory that embraced language families quite disparate from Indo-European, in the process correcting false formulations like Hungarian's being related to Slavic and other Indo-European languages. Like Columbus (who discovered the New World that large numbers of now-unknown scholars and explorers knew about, but who are never mentioned relative to this important historical event), and like Darwin (who integrated into one grand explanation of changes in organisms the isolated evolutionary facts discovered and known by various predecessors but who are generally unknown in this context, complementing the origin-of-species conception as mirrored in the six days enumerated in Genesis), Jones deserves the credit for revolutionizing thinking about linguistics. This great scholar, in one pithy, common-source formulation that may have been more quoted than any comparable formulation in any discipline excepting of course divine formulations like the Ten Commandments, gave ultimately the final, definitive rejection of the old mythological explanation of Hebrew given in the Garden of Eden to Adam as the first human language, and of God's destruction of the Tower of Babel as the explanation of language dispersion. This formulation ultimately separated language from religion and mythology, just as Buffon, Linneaus, and others had separated biota from the religious idea of forms unchanged since their creation. Jones so cleanly, delicately, and convincingly moved language study away from speculation, intuition, and religion, that contemporary theologians and others never realized what he had done to basic conceptions in Genesis, though it must be emphasized that he fervently believed in a Divine Being who loves all creatures. His formulation links human beings around the world within that love, of the sort that has motivated fine modern linguists like Kenneth Pike, and is the greatest linguistic formulation ever. Moreover, Jones's transmission of Sanskrit knowledge to the West and in a lesser sense his transmission of Persian and Arabic literature to Western civilization transcends this spectacular linguistic gift to posterity (see King 1990). It is comparable to Marco Polo's opening the magical world of Cathay to Europe (Cannon 1986: 186), even as the resulting cooperation with India would help lead to an Indian Renaissance and India's taking its place in the modern world as a leading country.

Dept. of English
Texas A&M University
College Station TX 77843-4227

ENDNOTES

Ernst Pulgram made many helpful suggestions in the revision of this paper.

REFERENCES

Cannon, Garland. 1986. "The construction of the European image of the Orient: A bicentenary reappraisal of Sir William Jones as poet and translator." *Comparative criticism* 8:166–88.

———. 1990. *The life and mind of Oriental Jones.* Cambridge: Cambridge University Press.

———. 1991. "Jones's 'sprung from some common source': 1786–1986." *Sprung from some common source.* Eds. Syndey M. Lamb & E. Douglas Mitchell. Stanford: Stanford University Press, pp. 23–47.

Dante. 1991. *De vulgari eloquentia: Dante's book of exile* Tr. Marianne Shapiro. Lincoln: University of Nebraska Press.

Emeneau, Murray. 1955. "India and linguistics." *Journal of the American Oriental Society* 75:145–53.

Jones, Sir William. 1807. *The works of Sir William Jones.* London: John Stockdale & John Walker. 13 vols.

———. 1970. *The letters of Sir William Jones.* Ed. Garland Cannon. Oxford: Clarendon Press. 2 vols.

King, Robert D. 1990. "West from India: The Odyssey of Sir William Jones." *Library chronicle* 20:49–63.

Mallory, James P. 1989. *In search of the Indo-Europeans.* London: Thames & Hudson.

ALLAN R. BOMHARD

The Nostratic Macrofamily (with special reference to Indo-European)

Abstract. Recent research makes it seem likely that Indo-European is not genetically isolated, but rather that it is distantly related to several other language families. This paper presents an overview of this research and then discusses how this research has thrown light on various problem areas within Indo-European.

From the very earliest days of Indo-European (IE) comparative linguistics, there have been speculations about the possible genetic relationship of IE to other language families. Though, in the course of study, many striking similarities were noted between IE and certain other language phyla, notably Uralic and Afroasiatic, truly convincing evidence of distant linguistic relationship was simply not brought forth. Gradually, the intellectual climate, especially in the United States, became hostile to long-range comparison, and IE remained an orphan with no known relatives.

Recently, within the past twenty years or so, the situation has slowly begun to turn around, and a growing number of linguists, especially in the former Soviet Union, have begun to turn attention toward investigating distant linguistic relationship. The new interest was sparked by the work of Vladislav Illič-Svityč and Aaron Dolgopolsky, who first started working independently and, at a later date, through the efforts of Vladimir Dybo, cooperatively. Their work, though not without its own shortcomings, was the first successful demonstration that certain language phyla of northern and central Eurasia, as well as the ancient Near East, might be genetically related. Following Pedersen, they employed the name "Nostratic" to designate this grouping of languages. In particular, Illič-Svityč, in the course of several publications, culminating in his posthumous comparative dictionary, which is still the process of publication, included IE, Kartvelian, Afroasiatic, Uralic, Dravidian, and Altaic in his version of the Nostratic macrofamily. From his very earliest writings, Dolgopolsky also included Chukchi-Kamchatkan.

At the present time, there is a handful of scholars, both within and outside the Soviet Union, devoting serious study to long-range comparison. Moreover, two organizations, namely, the Language Origins Society and the Association for the Study of Language in Prehistory, have been founded for the specific purpose of furthering the cause of investigating distant linguistic relationship. Some of the work being done is of very high quality, adhering quite strictly to the methodological principles established by the founders of IE comparative linguistics, while other work is quite speculative and less methodologically rigorous. Moreover, there are two main approaches being utilized: the first may be termed "taxonomy first", which seeks foremost to establish genetic relationship, while the second may be termed "reconstruction first", which emphasizes reconstruction. The first approach is reminiscent of the beginnings of IE comparative linguistics, where relationship was first established by the early pioneers such as Rasmus Rask, Franz Bopp, and Jacob Grimm, and it was only much later, beginning with August Schleicher, that actual reconstruction took place, though the need for reconstruction had been recognized as early as 1837 by Theodor Benfey. The two approaches are actually not mutually exclusive, but, rather, properly used, they can inform and further one another. I, personally, would give the edge to "taxonomy first". After all, one cannot successfully reconstruct until one has first established which languages might have a reasonable chance of being genetically related, that is to say, one must know which languages to compare.

The early founders of IE comparative linguistics placed great importance on the comparison of grammatical forms, and this bias continues to the present day in IE studies and has even been carried over into the study of other language phyla. However, this overemphasis on the comparison of grammatical forms is far too restrictive and was the reason that the Celtic languages, which have developed many unique features, were not immediately recognized as IE. Rather, as noted some sixty ago by Holger Pedersen:

> That agreement in the inflectional system is an especially clear and striking proof of kinship, no one denies. But it is only an anachronism in theory, which has no significance in actual practice, when such an agreement is still designated as the only valid proof. No one doubted, after the first communication about Tocharian . . . , that the language was Indo-European, though at that time virtually no similarities in inflection had been pointed out. Such similarities have since been shown, but even where they are almost obliterated, proof of kinship could be adduced from the vocabulary and from sound-laws. Hardly any one will assert that it would be impossible to recognize the relationship between, say, English and Italian, even without the help of other related languages

> or of older forms of these two languages themselves, although agreements between the inflectional systems are practically nonexistent.
>
> From the modern point of view it must be said that proof for relationship between languages is adduced by a systematic comparison of languages in their entirety, vocabulary as well as grammar. The reason why earlier scholars felt they should disregard the vocabulary was that they knew of no method of systematic comparison in this field (1931:245).

In 1957, Joseph Greenberg laid out a set of principles for establishing genetic relationship among languages, and these bear repeating. Greenberg notes that the simplest way to establish genetic relationship is by identifying a large number of similar morphs (or allomorphs)—especially irregularities—in similar environments in the languages being considered. Another significant indicator of probable genetic relationship is the presence of similar rules of combinability. Unfortunately, and this is significant, historical processes over the passage of time bring about the gradual transformation and eventual elimination of such similarities. The longer the period of separation, the less the chances will be that similarities of morphological forms and rules of combinability will be found. Fortunately, there are other factors that can be helpful in determining possible genetic relationship. One significant factor is the semantic resemblance of lexical forms. Here it is important to be able to establish recurrent sound-meaning correspondences for a reasonably large sample of lexical material. Lexical forms with identical or similar meanings have the greatest value. Next in value come forms that, though divergent in meaning, can convincingly be derived, through widely attested semantic shifts, from earlier forms with identical or similar meaning. The chances that lexical resemblances indicate genetic relationship increase dramatically when additional languages are brought into the comparison and when these new languages also exhibit a very large number of recurrent sound-meaning correspondences with the other languages. Greenberg considers the comparison of basic vocabulary from a large number of languages from a specific, wide geographic area to be the quickest and most reliable way to determine possible genetic relationship. To be meaningful, however, comparison must strive to eliminate chance resemblances and to separate borrowings from native elements. This is often easier said than done; however, Greenberg lays out two main techniques for detecting borrowed lexical items. First, he notes that borrowing is commonly confined to certain grammatical spheres (for example, cultural items) and certain grammatical categories (nouns far more often than verbs). Second, borrowed words can be distinguished from native vocabulary by expanding the range of comparison to include additional languages. It is only after these preliminary steps have been undertaken that meaningful comparison can begin. That is to say, and to reiterate, we must first have a good sense of which languages are likely candidates for comparison.

The basic principles underlying the Comparative Method may be summarized as follows: the first step involves the arduous task of data gathering. Once a large amount of lexical material has been gathered, it must be carefully analyzed to try to separate what is ancient from what is an innovation and from what is a borrowing. Once the native lexical elements have been reasonably identified within each phylum, the material can be compared across phyla to determine sound correspondences. Not only must the regular sound correspondences (that is, those that occur consistently and systematically) be defined, exceptions must also be noted and explained. Here, widely-attested sound changes (palatalization, metathesis, assimilation, dissimilation, syncope, etc.) provide the key to understanding the origin of most exceptions. In other cases, analysis of the influence that morphology has exerted will provide an understanding of how particular exceptions came into being. Some exceptions, however, though clearly related, simply defy explanation. All of these must be noted. The final step involves the reconstruction of the ancestral forms and the formulation of the sound laws leading to the forms in the descendant languages, identifying the laws that have produced the regular sound correspondences as well as the exceptions. The same principles apply to the reconstruction of the grammatical forms and rules of combinability and to the identification of the modifications leading to the systems found in the descendant languages. It was necessary to discuss these issues in order to address concerns that have been raised about the applicability of traditional methods of comparison to long-range comparison. It must be made perfectly clear that the same principles are just as applicable to long-range comparison as they are to any other type of linguistic comparison.

The first question that should be addressed is: what is the basis for setting up a Nostratic macrofamily? First and foremost, the descendant languages can be shown to share a large common vocabulary. In an article published in 1965, Illič-Svityč listed 607 possible common Nostratic roots, but only 378 have been published to date in his posthumous comparative Nostratic dictionary. Dolgopolsky currently claims to have approximately 1,500 common Nostratic roots, but none of this material has been published as yet. I have a great deal of lexical material (approximately 25,000 cited forms) from the Nostratic daughter languages to support 570 common Nostratic roots. This material will appear shortly in a joint monograph by myself and John C. Kerns, tentatively entitled *The Nostratic macrofamily*. It should be mentioned here as well that Greenberg is currently preparing a book entitled *Indo-European and its closest relatives: The Eurasiatic family*, in which a large amount of lexical material will be discussed. Notable among the lexical items uncovered by Illič-Svityč, Dolgopolsky, and myself is a solid core of common pronominal stems (these are listed in Table 1—only the stems represented in IE are given). As is to be expected, the various branches of Nostratic investigated to date exhibit regular sound correspondences (these are listed in Table 2), though, it should be mentioned, there are differences in interpretation between Illič-Svityč and Dolgopolsky on the one hand and myself on the other. Finally, a moderate number of common grammatical formants have been recovered.

The next question that needs to be answered is: which language phyla have been shown with probability to belong to the Nostratic macrofamily, and what subgroupings can be established? As noted earlier, Illič-Svityč included IE, Kartvelian, Afroasiatic, Uralic, Dravidian, and Altaic within the Nostratic macrofamily, and Dolgopolsky added Chukchi-Kamchatkan as well. Greenberg includes IE, Uralic-Yukaghir, Altaic (Mongolian, Chuvash-Turkic, and Manchu-Tungus), Japanese-Korean (Korean, Ainu, and Japanese-Ryukyuan), and Chukchi-Eskimo (Gilyak, Chukchi-Kamchatkan, and Eskimo-Aleut) in his Eurasiatic macrofamily. He does not include Kartvelian, Afroasiatic, nor Elamo-Dravidian—not because he believes that they are unrelated, but because he believes that these three language phyla are more distantly related to IE than are the others, which, along with IE, form a natural taxonomic subgrouping. My own opinion is close to that of Greenberg, though I would exclude Japanese-Ryukyuan and Ainu. We may note here that Paul Benedict has recently presented a large body of evidence to support his view that Japanese-Ryukyuan is genetically related to Formosan and other Austronesian languages, which is not to deny that there are Altaic elements in Japanese. Korean has, perhaps, the best chance of ultimately being an Altaic language, but much work still has to be done before this can be proved beyond a reasonable doubt. As I see the situation, Nostratic includes IE, Kartvelian, Afroasiatic, Uralic-Yukaghir, Elamo-Dravidian, Altaic, Chukchi-Kamchatkan, Gilyak (also called Nivkh), Eskimo-Aleut, and possibly Sumerian. Afroasiatic stands apart from the rest as an extremely an-

Table 1
The Distibution of Nostratic Pronoun Stems

Pronoun Stems	PIE	PK	PAA	PUr	PDr	PAlt	Sum
PERSONAL PRONOUN STEMS:							
1. **mi*/**me* 1 sg.: "I, me"	X	X	X	X		X	X
2. **ma*/**mə* 1 pl. incl.: "we, us"	X		X	X		X	X
3. **wa*/**wə* 1 pl.: "we, us"	X		X				
4. **na*/**nə* 1 pl.: "we, us"	X	X	X		X		
5. *$t[^h]i$/*$t[^h]e$ 2 sg.: "you	X		X	X		X	X
DEMONSTRATIVE STEMS:							
6. **sa*/**sə*	X	X		X			
7. *$t[^h]a$/*$t[^h]ə$ proximate	X		X	X	X	X	
8. *$t[^h]u$/*$t[^h]o$ distant	X		X	X			
9. *$k[^h]a$/*$k[^h]ə$	X	X	X				
10. *$d^y i$/$d^y e$	X		X	X			
11. **ʔi*/**ʔe* proximate	X	X		X	X	X	
12. **ʔa*/**ʔə* distant	X	X			X	X	
13. **na*/**nə*, **ni*/**ne*, **nu*/**no*	X		X	X			X
RELATIVE AND INTERROGATIVE STEMS:							
14. *$k^w[^h]i$/*$k^w[^h]e$ relative	X			X		X	
15. *$k^w[^h]a$/*$k^w[^h]ə$ interrogative	X		X	X		X	
16. **mi*/**me* interrogative	X	X	X	X		X	X
17. **ma*/**mə* relative	X	X	X	X			
18. **ʔay* interrogative & relative	X		X	X	X	X	

EXAMPLES (PERSONAL PRONOUNS)

1 and 2. Proto-Nostratic **mi-*/**me-* 1st sg. personal pronoun stem: 'I, me'; **ma-*/**mə-* inclusive 1st pl. personal pronoun stem: 'we, us':

A. PIE **me-*/**mo-* 1st person personal pronoun stem: Sanskrit (acc. sg.) *mā, mā́m;* Greek (acc. sg.) μɛ, ἐμέ; Latin (acc.-abl. sg.) *mē;* Old Irish (nom. sg.) *mé, messe* 'I', (acc. sg.) *mé, messe, -m* 'me'; Gothic (acc. sg.) *mik;* Lithuanian (acc. sg.) *manę̃*; Old Church Slavic (acc. sg.) *mę, mene;* Hittite *-mi, -mu; am-mu-uk, mi-iš.* Verb endings: Sanskrit (1st sg. primary active) *-mi,* (1st pl. primary active) *-mas,* (1st pl. primary middle; 1st pl. secondary perfect) *-mahe,* (1st pl. secondary active; 1st pl. secondary perfect) *-ma,* (1st pl. secondary middle) *-mahi*; Hittite (1st sg. active *mi*-conjugation) *-mi,* (1st pl. active *mi*-conjugation; if the stem ends in *-u-*) *-meni, -mani,* (1st pl. preterite *mi*-conjugation; if the stem ends in *-u-*) *-men;* Greek (Homeric) (1st sg. active indicative athematic) -μι, (1st pl. active indicative) -μεν, (1st sg. middle indicative) -μαι, (1st dual middle indicative) -μεθον, (1st pl. middle indicative) -μεθα (-μεσθα).

B. Proto-Kartvelian **me-, *men-* 1st person personal pronoun stem: Georgian *me-, men-, mena-;* Mingrelian *ma-;* Zan *ma, man;* Svan *mi-*. Georgian *m-* 1st person sg. verb prefix (objective conjugation).

C. Proto-Afroasiatic **ma-/*mə-* 1st person personal pronoun stem: Chadic, Hausa (pl.) *ma* 'we', (indirect object pl.) *manà* 'us, to us, for us', (pl.) *mu* 'we, us, our', (past tense subj. pl.) *mun* 'we', (continuous tense subj. pl.) *munà* 'we'; (indirect object sg.) *minì* 'me, to me, for me'; Kotoko *mi* 'we, us'; Mandara *ma* 'we, us'; Musgu *mi* 'we, us', *mu* 'I, me'; Bole *mu* 'we, us'.

D. Proto-Uralic **me* 1st person sg. personal pronoun stem: 'I, me': Finnish *minä/minu-* 'I, me'; Lapp *mon/mú-;* Mordvin *mon;* Cheremis / Mari *mən;* Votyak / Udmurt *mon;* Zyrian / Komi *me,* (acc.) *menõ;* Ostyak / Xanty *mä/mən-*, (N.) *ma/măn-*, (S.) *mæn/mən-;* Yurak Samoyed / Nenets *mań;* Tavgi Samoyed / Nganasan *mannaŋ;* Yenisei Samoyed / Enets *mod'i;* Selkup Samoyed *man, mat;* Kamassian *man*. Yukaghir *met*. Proto-Uralic **me* 1st person pl. personal pronoun stem: 'we': Finnish *me* 'we'; Lapp *mí;* Mordvin *min;* Cheremis / Mari *mä, me;* Votyak / Udmurt *mi;* Zyrian / Komi *mi;* Vogul / Mansi *man;* Ostyak / Xanty *mŏŋ;* Hungarian *mi;* Yurak Samoyed / Nenets *mańaˀ*; Tavgi Samoyed / Nganasan *meeŋ;* Yenisei Samoyed / Enets *mod'iˀ*; Selkup Samoyed *mee, mii;* Kamassian *miˀ*. Yukaghir *mit*. Proto-Uralic 1st person personal suffix **-me*.

E. Altaic: Classical Mongolian (nom. sg.) *bi* 'I', (gen. sg.) *minu;* (nom. pl. exclusive) *ba* 'we' (= 'I and the ones about me') (the nom. pl. inclusive is *bide* 'you and I'), (gen. pl.) *manu,* (acc. pl). *mani*; Old Turkish (nom. sg.) *män* (rarely *bän*), (gen. sg.) *mäniŋ* (rarely *bäniŋ*), (dat. sg.) *maŋa* (*baŋa*), (acc. sg.) *mäni* (*mini, bini*), (loc. sg.) *mintä, mintäda, mindidä,* (abl. sg.) *minitin;* Yakut (acc. sg.) *mana,* (gen. sg.) *man*(*ăn*), (loc. sg.) *manra,* (abl. sg.) *manran,* (causalis) *manšăn.*

F. Sumerian (Emesal) *ma(-e), me-a, me-e* 'I', (1st pl. possessive suffix) *-me* 'our'.

3. Proto-Nostratic **wa-/*wə-* 1st person personal pronoun stem:

A. PIE **we-/*wo-, *wey-* 1st person dual and plural personal pronoun stem: Hittite *ú-i-e-eš* 'we'; Sanskrit (dual) *vā́m,* (pl.) *vayám* 'we'; Avestan (dual) *vā,* (pl.) *vaēm* 'we'; Gothic (dual) *wit,* (pl.) *weis* 'we'; Old Church Slavic (dual) *vě* 'we two'; Tocharian A *was,* B *wes* 'we'. Verb endings: Luwian (1st sg.) *-wi;* Hittite (1st pl.) *-weni, -wani, -wen, -wašta, -waštati, -waštat;* Sanskrit (1st dual) *-vas, -va, -vahe, -vahi.*

B. Proto-Afroasiatic **wa-/*wə-* 1st person personal pronoun stem: Egyptian *wy* 'I, me'. Chadic: Ngizim *wà* (inclusive) 'we, us, our'.

4. Proto-Nostratic **na-/*nə-* 1st person personal pronoun stem:

A. PIE **ne-/*no-/*n̥-s-* used in the oblique cases of the personal pronoun of the first person dual and plural: Sanskrit (dual) *nau,* (pl.) *nas* 'us'; Greek (dual) νώ 'us'; Latin *nōs* 'we, us'; Gothic *uns* 'us'; Old Church Slavic *nasъ* 'us'; Hittite *an-za-a-aš* 'us'.

B. Kartvelian: Svan *näj* 'we'.

C. Proto-Afroasiatic _*na-/*nə-_ 1st person personal pronoun stem: Proto-Semitic _*-nā̆_ 1st pl. personal pronoun suffix, _*na-/*ni-_ 1st pl. personal pronoun prefix > Hebrew _-nū_, _ni-_; Aramaic _-n(ā)_, _ne-;_ Ugaritic _-n_, _n-_ ; Akkadian _-āni_, _-ānu; ni-;_ Arabic _-nā_, _na-;_ Geez / Ethiopic _-na_, _nə-_. Proto-Semitic (independent 1st pl. personal pronoun) _*naħnū̆ 'we'_ > Hebrew _(ʔă)naḥnū;_ Akkadian _nīnu;_ Arabic _naḥnu;_ Jibbāli _nḥán;_ Ḥarsūsi _neḥā;_ Mehri _nəḥā n-;_ Geez / Ethiopic _nəḥna;_ Tigrinya _nəḥna_. Egyptian _ínn_ 'we', (suffix) _-n;_ Coptic _anon_, _an-_, _ann-_ 'we'. Berber: Tamazight (1st sg. independent personal pronoun) _nəkk_ 'I', (1st pl. independent personal pronoun) _nukni_ 'we', (1st pl. indirect pronoun) _nəx_ (possessive _-nnəx_). Proto-East Cushitic _*na/*ni/*nu_ 'we' > Burji _náanu_ 'we', _nín-ka_ 'our', _nín-si_ 'us'; Saho _nanu_ 'we'; Oromo _nu_ 'we'; Gedeo (nom. pl.) _noʔ_o 'we', (acc. pl.) _noʔo(o)_ 'us', (dat. pl.) _noʔoʔá_, _noʔá_ 'to us', (poss.) (m.) _noʔo-ka_, (f.) _noʔo-tt'a_ 'our'; Hadiyya (nom. pl.) _neese_ 'we', (acc. pl.) _ne(e)s_ 'us', (dat. pl.) _niin_ 'to us', (poss.) _ni-_ 'our'; Kambata (nom. pl.) _naʔooti_ 'we', (acc. pl.) _ne(e)s_, _-nne_ 'us', (dat. pl.) _nesá_ 'to us', (poss.) _-nne_ 'our'; Sidamo (nom. -acc. pl.) _ninke_ 'we, us', (dat. pl.) _ninke-ra_ 'to us', (poss.) _-nke_ 'our'. Proto-Southern Cushitic _*nana_, _*nani_ 'we' > Ma'a _níne_ 'we'; Dahalo _nányi_, _nyányi_ 'we'. Chadic: Hausa _ni_ 'I, me'; Ngizim _na(a)_ 'I'.

D. Proto-Dravidian _*nām-_, _*nam-_ 'we' (inclusive): Tamil _nām_ (obl. _nam(m)-_) 'we'; Malayalam _nām_ (obl. _nam(m)-_) 'we'; Kannaḍa _nāvu_ (obl. _nam-_) 'we'; Tuḷu _nama_ 'we'; Kuṛux _nām (man-, naŋg-)_ 'we'; Malto _nám_ (obl. _nam-_) 'we'; Brahui _nan_ 'we'.

5. Proto-Nostratic _*t[^{h}]i/*t[^{h}]e_ 'you':

A. PIE (nom. sg.) _*t[^{h}]ū̆_ 'you'; (acc. sg.) _*t[^{h}]wē̆/*t[^{h}]ē̆_, _*t[^{h}]wēm/*t[^{h}]ēm:_ Sanskrit (nom. sg.) _tvám_ 'you', (acc. sg.) _tvā́m_, _tvā;_ Avestan (nom. sg.) _tūm_, _tū_ 'you'; Greek (Doric) (nom. sg.) _tύ_ 'you', (acc. sg.) _tέ_; Armenian (nom. sg.) _du_ 'you'; Albanian (nom. sg.) _ti_ 'you', (acc. sg.) _ty_, _të;_ Latin (nom. sg.) _tū_ 'you', (acc. sg.) _tē;_ Old Irish (nom. sg.) _tú_ 'you'; Gothic (nom. sg.) _þu_ 'you', (acc. sg.) _þuk;_ Lithuanian (nom. sg.) _tù_ 'you', (acc. sg.) _tavę́;_ Old Church Slavic (nom. sg.) _ty_ 'you', (acc. sg.) _tę_, tebe; Palaic (nom. sg.) _ti-i_ 'you', (dat.-acc. sg.) _tu-ú;_ Hittite (nom. sg.) _zi-ik_, _zi-ga_ 'you', (acc.-dat. sg.) _tu-uk_, _tu-ga_, (gen. sg.) _tu-(e-)el_, (abl. sg.) _tu-e-da-az_, _tu-e-ta-za;_ (encl. possessive nom. sg.) _-ti-iš_, (encl. possessive acc. sg.) _-ti-in_, (encl. possessive neut. sg.) _-te-it_, (encl. possessive gen. sg.) _-ta-aš_, (encl. possessive dat. sg.) _-ti_, (encl. possessive instr. sg.) _-te-it;_ (encl. oblique sg.) _-ta_ (_-du_ before _-za_). Note: the PIE reconstructions represent later, post-Anatolian forms. PIE (2nd pl. verb ending) _*-t[^{h}]e:_ Sanskrit (2nd pl. primary verb ending) - _tha_, (2nd pl. secondary verb ending) _-ta;_ Greek _-tε;_ Latin (imptv.) _-te;_ Old Irish _-the_, _-de;_ Gothic _-þ;_ Lithuanian _-te;_ Old Church Slavic _-te_.

B. Proto-Afroasiatic _*t[^{h}]a/*t[^{h}]ə_ 'you': Proto-Semitic (prefix forms) _*t[^{h}]i-/*t[^{h}]a-_, (suffix forms) _*-t[^{h}]ī̆/-*t[^{h}]ā̆_ 'you'; Arabic (m.) _ʔan-ta_, (f.) _ʔan-ti_ 'you'; perfect 2nd sg. endings (m.) _-ta_, (f.) _-ti;_ imperfect 2nd sg./du./pl prefix _ta-;_ Akkadian (m.) _at-tā_, (f.) _at-tī_ 'you'; permansive 2nd sg. endings (m.) _-āt(a)_, (f.) _-āti;_ prefix conjugation 2nd sg./pl. prefix _ta-;_ Hebrew (m.) _ʔat-tāh_, (f.) _ʔat-t(ī)_ 'you'; perfect 2nd sg. endings (m.)

-tā, (f.) *-t(i);* imperfect 2nd sg./pl. prefix *ti-;* Ugaritic *át* 'you' (m. **ʔatta*, f. **ʔatti*); perfect 2nd sg. ending *-t* (m. **-ta*, f. **-ti*); imperfect 2nd sg./du./pl. prefix *t-;* Jibbāli *tɛn* 'you'; Geez / Ethiopic (m.) *ʔan-ta*, (f.) *ʔan-tī* 'you'; prefix conjugation 2nd sg./pl. prefix *tə-*. Cushitic: Bilin (sg.) *en-ti, in-ti* 'you', (pl.) *en-tin, in-tin*. Proto-East Cushitic **ʔ̱at-i/u* (2nd sg. subj.) 'you' > Oromo *at-i* 'you'; Gedeo *at-i* 'you'; Hadiyya *at-i* 'you'; Kambata *at-i* 'you'; Sidamo *at-e/i* 'you'; Burji *áši* 'you'; Somali *ad-i-* 'you'; Saho-Afar *at-u* 'you'; Rendille *at-i* 'you'; Bayso *at-i* 'you'; Konso *at-ti* 'you'; Gidole *at-te* 'you'. Proto-East Cushitic *ʔ̱atin- (2nd pl. subj.) 'you' > Saho-Afar *atin* 'you'; Burji *ašinu* 'you'; Somali *idin-* 'you'; Rendille *atin-* 'you'; Dasenech *itti(ni)* 'you'; Kambata *a*ʔn-a*ʔooti* 'you'; Tsamay *atun-i* 'you'. Proto-Highland East Cushitic **tee* (2nd sg. voc. fem.) 'you' > Gedeo *tee* (f.) 'you'; Hadiyya *ta* (f.) 'you'; Kambata *te* (f.) 'you'; Sidamo *tee* (f.) 'you'. Proto-Southern Cushitic **ʔ̱ata-* (plural) 'you', **ʔ̱aata-* (sing.) 'you'; Iraqw *aten* 'we'; Dahalo *ʔāṯṯa* (plural) 'you', *ʔááṯa* (sing.) 'you'.

C. Proto-Uralic (sg.) **te* 'you': Finnish *sinä/sinu-* 'you'; Lapp *don/dú-* 'you'; Mordvin *ton* 'you'; Cheremis / Mari *təń* 'you'; Votyak / Udmurt *ton* 'you'; Zyrian / Komi *te* (acc. *tenõ*) 'you'; Hungarian *të* 'you'; Tavgi Samoyed / Nganasan *tannaŋ* 'you'; Yenisei Samoyed / Enets *tod'i* 'you'; Selkup Samoyed *taŋ, tat* 'you'; Kamassian *tan* 'you'. Yukaghir *tet* 'you'. Proto-Uralic (pl.) **te* 'you': Finnish *te* 'you'; Lapp *dí* 'you'; Mordvin (Erza) *tiń, tïń* 'you'; Cheremis / Mari *tä, te* 'you'; Votyak / Udmurt *ti* 'you'; Zyrian / Komi *ti* 'you'; Hungarian *ti* 'you'; Tavgi Samoyed / Nganasan *teeŋ* 'you'; Yenisei Samoyed / Enets *tod'iʔ* 'you'; Selkup Samoyed *tee, tii* 'you'; Kamassian *šiʔ* 'you'. Yukaghir *tit* 'you'.

D. Elamo-Dravidian: Elamite (2nd sg. verb ending) *-t*, (2nd pl. verb ending) *-ht* (*h* + *t;* in Royal Achaemenid Elamite, this becomes *-t* due to loss of *h*), allocutive (that is, person addressed or 'second person') gender suffix *-t*. Dravidian: Parji *-t* appositional marker of 2nd sg. in pronominalized nouns and verb suffix of 2nd sg.

E. Altaic: Mongolian (nom. sg.) *(*ti >) či* 'you', (nom. pl.) *ta;* Turkish (nom. sg.) *sen* 'you' (acc. sg. *seni*, gen. sg. *senin*, dat. sg. *sana*), (nom. pl.) *siz* 'you' (acc. pl. *sizi*, gen. pl. *sizin*, dat. pl. *size*); Chuvash (nom. sg.) *esĕ* 'you', (obl.) *san;* (nom. pl.) *esir*, (obl. pl.) *sir* (the prefixed vowel *e-* in the nominative singular and plural is a particle; the original forms are **si* (obl. **sän*) and **sir*).

F. Sumerian *za-e* 'you', (2nd sg. possessive suffix) *-zu* 'your'.

cient, independent branch. Younger are Kartvelian and Elamo-Dravidian. IE, Uralic-Yukaghir, Altaic, Gilyak, Chukchi-Kamchatkan, and Eskimo-Aleut appear to be more closely related as a group than any one of them is to Afroasiatic, Kartvelian, and Elamo-Dravidian. Finally, Sumerian, if it really does belong here, is a separate branch, probably closest to Elamo-Dravidian. To be sure, there remain numerous problems to be resolved, such as, for example, whether or not

TABLE 2: NOSTRATIC SOUND CORRESPONDENCES

Proto-Nostr.	Proto-IE	Proto-Kartv.	Proto-AA	Proto-Uralic	Proto-Dravidian	Proto-Altaic	Sum.
b-	*b*[^{h}]-	*b-*	*b-*	*p-*	*p-*	*b-*	*b-*
-b-	-*b*[^{h}]-	*-b-*	*-b-*	*-w-*	*-pp-/-vv-*	*-b-*	*-b-*
p[^{h}]-	*p*[^{h}]-	*p*[^{h}]-	*p*[^{h}]-	*p-*	*p-*	*p-*	*p-*
-*p*[^{h}]-	-*p*[^{h}]-	-*p*[^{h}]-	-*p*[^{h}]-	*-p-*	*-pp-/-v-*	*-p-/-b-*	*-p-*
d-	*d*[^{h}]-	*d-*	*d-*	*t-*	*t-*	*d-*	*d-*
-d-	-*d*[^{h}]-	*-d-*	*-d-*	*-δ-*	*-ṭ(ṭ)-*	*-d-*	*-d-*
t[^{h}]-	*t*[^{h}]-	*t*[^{h}]-	*t*[^{h}]-	*t-*	*t-*	*t-*	*t-*
-*t*[^{h}]-	-*t*[^{h}]-	-*t*[^{h}]-	-*t*[^{h}]-	*-t(t)-*	*-t(t)-*	*-t-*	*-t-*
t'-	*t'-*	*t'-*	*t'-*	*t-*	*t-*	*t-*	*d-*
-t'-	*-t'-*	*-t'-*	*-t'-*	*-t-*	*-t(t)-*	*-d-*	*-d-*
d^{y}-	*d*[^{h}]-	*d*$^{\check{z}}$-	*d*^{y}-	*t*^{y}-	*c-*	*d*$^{\check{z}}$-	*d-*
-*d*^{y}-	-*d*[^{h}]-	-*d*$^{\check{z}}$-	-*d*^{y}-	-*t*^{y}-	*-c(c)-*	-*d*$^{\check{z}}$-/-*d*-	*-d-*
t^{y}[^{h}]-	*t*[^{h}]-	*t*$^{\check{s}}$[^{h}]-	*t*^{y}[^{h}]-	*t*^{y}-	*c-*	*t*$^{\check{s}}$-	*š-*
-*t*^{y}[^{h}]-	-*t*[^{h}]-	-*t*$^{\check{s}}$[^{h}]-	-*t*^{y}[^{h}]-	-*t*^{y}-	*-c(c)-*	-*t*$^{\check{s}}$-	*-š-*
t'y-	*t'-*	*t*$^{\check{s}}$'-	*t'y-*	*t*^{y}-	*c-*	*t*$^{\check{s}}$-	*d-*
-t'y-	*-t'-*	-*t*$^{\check{s}}$'-	*-t'y-*	-*t*^{y}*t*^{y}-	*-c(c)-*	-*d*$^{\check{z}}$-	*-d-*

Proto-Nostr.	Proto-IE	Proto-Kartv.	Proto-AA	Proto-Uralic	Proto-Dravidian	Proto-Altaic	Sum.
s^{y}–	s–	š–	s^{y}–	s^{y}–	c–	s–	š–
–s^{y}–	–s–	–š–	–s^{y}–	–s^{y}–	–c(c)–/–y–	–s–	–š–
d^{z}–	$d[^{h}]$–	$d^{z}{}_{1}$–	d^{z}–	$t^{\check{s}}$–	c–	$d^{\check{z}}$–	z–
–d^{z}–	–$d[^{h}]$–	–$d^{z}{}_{1}$–	–d^{z}–	–$t^{\check{s}}$–	–c(c)–	–$d^{\check{z}}$–/–d–	–z–
$t^{s}[^{h}]$–	$t[^{h}]$–	$t^{s}[^{h}]_{1}$–	$t^{s}[^{h}]$–	$t^{\check{s}}$–	c–	$t^{\check{s}}$–	s–
–$t^{s}[^{h}]$–	–$t[^{h}]$–	–$t^{s}[^{h}]_{1}$–	–$t^{s}[^{h}]$–	–$t^{\check{s}}$–	–c(c)–	–$t^{\check{s}}$–	–s–
$t^{s'}$–	t'–	$t^{s'}{}_{1}$–	$t^{s'}$–	$t^{\check{s}}$–	c–	$t^{\check{s}}$–	z–
–$t^{s'}$–	–t'–	–$t^{s'}{}_{1}$–	–$t^{s'}$–	–$t^{\check{s}}$–	–c(c)–	–$d^{\check{z}}$–	–z–
s–	s–	s_{1}–	s–	s–	c–	s–	s–
–s–	–s–	–s_{1}–	–s–	–s–	–c(c)–	–s–	–s–
$t^{l}[^{h}]$–	$k[^{h}]$–	x–	$t^{l}[^{h}]$–	s^{y}–	c–	k–	
–$t^{l}[^{h}]$–	–$k[^{h}]$–	–x–	–$t^{l}[^{h}]$–	–s^{y}–	–k–	–k–/–g–	
$t^{l'}$–	k'–		$t^{l'}$–	δ^{y}–	t–	k–	d–
–$t^{l'}$–	–k'–		–$t^{l'}$–	–δ^{y}–	–ṭ(ṭ)–	–g–	–d–

Proto- Nostr.	Proto- IE	Proto- Kartv.	Proto- AA	Proto- Uralic	Proto- Dravidian	Proto- Altaic	Sum.
g-	$g[^{h}]$-	*g-*	*g-*	*k-*	*k-*	*g-*	*g-*
-g-	-$g[^{h}]$-	*-g-*	*-g-*	*-γ-*	*-k-*	*-g-*	*-g-*
$k[^{h}]$-	$k[^{h}]$-	$k[^{h}]$-	$k[^{h}]$-	*k-*	*k-*	*k-*	*k-*
-$k[^{h}]$-	-$k[^{h}]$-	-$k[^{h}]$-	-$k[^{h}]$-	*-k(k)-*	*-k(k)-*	*-k-/-g-*	*-k-*
k'-	*k'-*	*k'-*	*k'-*	*k-*	*k-*	*k-*	*g-*
-k'-	*-k'-*	*-k'-*	*-k'-*	*-k-*	*-k(k)-*	*-g-*	*-g-*
g^{y}-	$g[^{h}]$-	*g-*	g^{y}-	*k-*	*k-*	*g-*	*g-*
-g^{y}-	-$g[^{h}]$-	*-g-*	-g^{y}-	*-γ-*	*-k-*	*-g-*	*-g-*
$k^{y}[^{h}]$-	$k[^{h}]$-	$k[^{h}]$-	$k^{y}[^{h}]$-	*k-*	*k-*	*k-*	*k-*
-$k^{y}[^{h}]$-	-$k[^{h}]$-	-$k[^{h}]$-	-$k^{y}[^{h}]$-	*-k(k)-*	*-k(k)-*	*-k-/-g-*	*-k-*
k'^{y}-	*k'-*	*k'-*	k'^{y}-	*k-*	*k-*	*k-*	*g-*
-k'^{y}-	*-k'-*	*-k'-*	-k'^{y}-	*-k-*	*-k(k)-*	*-g-*	*-g-*
g^{w}-	$g^{w}[^{h}]$-	*gw/u-*	g^{w}-	*k-*	*k-*	*g-*	*gu-*
-g^{w}-	-$g^{w}[^{h}]$-	*-gw/u-*	-g^{w}-	*-γ-*	*-k-*	*-g-*	*-gu-*
$k^{w}[^{h}]$-	$k^{w}[^{h}]$-	$k[^{h}]w/u$-	$k^{w}[^{h}]$-	*k-*	*k-*	*k-*	*ku-*
-$k^{w}[^{h}]$-	-$k^{w}[^{h}]$-	-$k[^{h}]w/u$-	-$k^{w}[^{h}]$-	*-k(k)-*	*-k(k)-*	*-k-/-g-*	*-ku-*
k'^{w}-	k'^{w}-	*k'w/u-*	k'^{w}-	*k-*	*k-*	*k-*	*gu-*
-k'^{w}-	-k'^{w}-	*-k'w/u-*	-k'^{w}-	*-k-*	*-k(k)-*	*-g-*	*-gu-*

Proto- Nostr.	Proto- IE	Proto- Kartv.	Proto- AA	Proto- Uralic	Proto- Dravidian	Proto- Altaic	Sum.
ɢ-	g[ʰ]-	ɢ-	g-	k-	k-	g-	g-
-ɢ-	-g[ʰ]-	-ɢ-	-g-	-γ-	-k-	-g-	-g-
q'-	k'-	q'-	k'-	k-	k-	k-	g-
-q'-	-k'-	-q'-	-k'-	-k-	-k(k)-	-g-	-g-
q'ʷ-	k'ʷ-	q'w/u-	k'ʷ-	k-	k-	k-	gu-
-q'ʷ-	-k'ʷ-	-q'w/u-	-k'ʷ-	-k-	-k(k)-	-g-	-gu-
ʕ-	ʕ͜ɦ-	Ø-	ʕ-	Ø-	Ø-	Ø-	
-ʕ-	-ʕ͜ɦ-	-Ø-	-ʕ-	-Ø-	-Ø-	-Ø-	
ħ-	ħ͜h-	x-	ħ-	Ø-	Ø-	Ø-	h-
-ħ-	-ħ͜h-	-x-	-ħ-	-Ø-	-Ø-	-Ø-	-h-
ʔ-	ʔ-	Ø-	ʔ-	Ø-	Ø-	Ø-	Ø-
-ʔ-	-ʔ-	-Ø-	-ʔ-	-Ø-	-Ø-	-Ø-	-Ø-
h-	h-	Ø-	h-	Ø-	Ø-	Ø-	Ø-
-h-	-h-	-Ø-	-h-	-Ø-	-Ø-	-Ø-	-Ø-
y-	y-	y-/Ø-	y-	y-	y-/Ø-	y-	
-y-	-y-		-y-	-y-	-y-	-y-	
w-	w-	w-	w-	w-	v-/Ø-		
-w-	-w-	-w-	-w-	-w-	-v-		

Proto-Nostr.	Proto-IE	Proto-Kartv.	Proto-AA	Proto-Uralic	Proto-Dravidian	Proto-Altaic	Sum.
m-	*m-*	*m-*	*m-*	*m-*	*m-*	*m-*	*m-*
-m-	*-m-*	*-m-*	*-m-*	*-m-*	*-m-*	*-m-*	*-m-*
n-	*n-*		*n-*	*n-*	*n-*		*n-*
-n-	*-n-*	*-n-*	*-n-*	*-n-*	*-n-/-ṉ-*	*-n-*	*-n-*
n^{y}-	*n-*		*n-*	n^{y}-		n^{y}-	
-n^{y}-	*-n-*		*-n-*	-n^{y}-	*-ṇ-*	-n^{y}-	
-ŋ-	*-n-*		*-n-*	*-ŋ-*	*-ṇ-*	*-ŋ-*	
l-	*l-*	*l-*	*l-*	*l-*	*l-*		*l-*
-l-	*-l-*	*-l-*	*-l-*	*-l-*	*-l-*	*-l-*	*-l-*
r-	*r-*	*r-*	*r-*	*r-*			*r-*
-r-	*-r-*	*-r-*	*-r-*	*-r-*	*-r-/-ṟ-*	*-r-*	*-r-*

Proto-Nostratic	Proto-Indo-European	Proto-Kartvelian	Proto-Afroasiatic
i	*i, e*	*i*	*ə*
ə	*e, a, ə*	*e, i*	*ə*
u	*u, o*	*u*	*ə*
e	*e*	*e*	*a*
a	*a, o, ə*	*a*	*a*
o	*o*	*o*	*a*
iy	*ĭy, ey; ī, ē; ĭ*	*iy, i*	*əy*
əy	*ey, ay, ĭy, ĭ*	*ey, i*	*əy*
uy	*ĭy; ī; ĭ*	*uy, i*	*əy*
ey	*ey, ĭy; ē; ĭ*	*ey, i*	*ay*
ay	*ay, oy, ĭy, ĭ*	*ay, i*	*ay*
oy	*oy, ĭy, ĭ*	*oy, i*	*ay*
iw	*ū, uw, ŭ*	*iw, u*	*əw*
əw	*ew, aw, ŭw, ŭ*	*ew, u*	*əw*
uw	*ū, ō, ŭw, ow, ŭ*	*uw, u*	*əw*
ew	*ew, ŭw, ŭ*	*ew, u*	*aw*
aw	*ow, ŭw, ŭ*	*aw, u*	*aw*
ow	*ō, ow, ŭw, ŭ*	*ow, u*	*aw*

Proto-Nostratic	Proto-Uralic	Proto-Dravidian	Proto-Altaic	Sumerian
i	*i*	*i*	*i, ï*	*i*
ə	*e*	*e*	*e*	*e*
u	*u*	*u*	*u, ü*	*u*
e	*e*	*e*	*e*	*e*
a	*a, ä*	*a*	*a*	*a*
o	*o*	*o*	*o, ö*	*u*
iy	*iy, i*	*iy, ī*	*ī, ï̄*	*i*
əy	*ey*	*ey, ī*	*ē; i, ï*	*i*
uy	*uy*	*uy, ī*		*i*
ey	*ey, e*	*ey, ē*	*ėy; ė̄*	*e*
ay	*ay, äy*	*ay, ā*	*a; i, ï*	*e*
oy	*oy*	*oy, ō*		*e*
iw	*iw*	*iv, ū*		*u*
əw	*ew*	*ev, ū*		*u*
uw	*uw, u*	*uv, ū*	*ū, ǖ*	*u*
ew	*ew*	*ev, ē*		*u*
aw	*aw, äw*	*av, ā*	*ō, ȫ*	*u*
ow	*ow, o*	*ov, ō*	*ō, ȫ*	*u*

Altaic is even a valid taxonomic entity, but, in due course, as more and more scholars turn their attention to these issues, solutions will begin to emerge.

Let us now look at IE and address the question: what is gained by comparing IE with the other Nostratic languages? The following gains may be mentioned as being among the most important: (a) a better understanding of the laryngeals, (b) a better understanding of root structure patterning, (c) a better understanding of the origin of verb morphology, (d) a better understanding of the origin and development of vowel gradation, and (e) support for the Hopper, Gamkrelidze, and Ivanov reinterpretation of IE consonantism. We may now look at each one of these in more detail.

A. Laryngeals. According to Kuryłowicz and those who follow his theories (such as Sturtevant and Lehmann, among others), IE is assumed to have had four laryngeals, which may be symbolized as $*H_1$, $*H_2$, $*H_3$, and $*H_4$. Other scholars posit only three laryngeals, denying the existence of $*H_4$, and, still others, posit as few as one laryngeal or as many as twelve. We will stick with the four laryngeals posited by Kuryłowicz. Now, of the other Nostratic branches, only Afroasiatic has a full set of laryngeals. Though Semitic is traditionally assumed to have had six laryngeals, the Afroasiatic parent language most likely had only four, namely, a glottal stop, a voiceless laryngeal (or glottal) fricative, and voiced and voiceless pharyngeal fricatives. Extremely good correspondences can be established between Afroasiatic and IE, and, as a result, it is now possible to establish the probable phonetic values of the laryngeals in IE. We can confirm that $*H_1$ was a glottal stop and $*H_4$ was a voiceless laryngeal fricative as originally suggested by Sapir, Sturtevant, and Lehmann, while $*H_2$ was probably the voiceless and voiced multiply-articulated pharyngeal/laryngeal fricatives, and $*H_3$ was probably originally identical to $*H_2$. That is to say, there is no evidence from the cognate Nostratic languages to support positing $*H_3$ distinct from $*H_2$. Note that both of these two laryngeals have the same reflex in Hittite. The only reason that two separate laryngeals were set up in IE in the first place was to account for several cases of nonapophonic *o*. However, these examples can be accounted for much better by assuming that this single, combined $*H_2$ and $*H_3$ changed a contiguous original *u* to *o* along the lines of what is found in modern Arabic dialects.

B. Root structure patterning. Comparison of IE with the other Nostratic branches, especially Kartvelian and Afroasiatic, allows us to refine the theories of Benveniste and, in so doing, to trace the development of that patterning from the earliest times down to the appearance of the individual daughter languages. The most ancient patterning may be assumed to have been as follows:

1. There were no initial vowels in the earliest form of pre-IE. Therefore, every root began with a consonant.
2. Originally, there were no initial consonant clusters either. Consequently, every root began with one and only one consonant.
3. Two basic syllable types existed: (A) **CV* and (B) **CVC*, where *C* = any nonsyllabic and *V* = any vowel. Permissible root forms coincided exactly with these two syllable types.
4. A verbal stem could either be identical with a root or it could consist of a root plus a single derivational morpheme added as a suffix to the root: **CVC-VC-*. Any consonant could serve as a suffix.
5. Nominal stems, on the other hand, could be further extended by additional suffixes.

In the earliest form of IE, there were three fundamental stem types: (a) verbal stems, (b) nominal and adjectival stems, and (c) pronominal and indeclinable stems.

The phonemicization of a strong stress accent disrupted the patterning outlined above. The positioning of the stress was morphologically distinctive, serving as a means to differentiate grammatical categories. All vowels were retained when stressed but were either weakened (= 'reduced-grade') or totally eliminated (= 'zero-grade') when unstressed: the choice between the reduced-grade versus the zero-grade depended upon the position of the unstressed syllable relative to the stressed syllable as well as upon the laws of syllabicity in effect at that time. Finally, it was at this stage of development that the syllabic allophones of the resonants came into being.

The stress-conditioned ablaut alternations gave rise to two distinct forms of extended stems:

Type 1: root in full-grade and accented, suffix in zero-grade: **CVCC-*.
Type 2: root in zero-grade, suffix in full-grade and accented: **CCVC-*.

When used as a verbal stem, Type 1 could undergo no futher extension. However, Type 2 could be further extended by means of a 'determinative'. Further addition of a determinative or suffixes pointed to a nominal stem. According to Benveniste, a 'suffix' was characterized by two alternating forms (**-et-/*-t-*, **-en-/*-n-*, **-ek-/*-k-*, etc), while a 'determinative' was characterized by a fixed consonantal form (**-t-*, **-n-*, **-k-*, etc). In its beginnings, ablaut was merely a phonological alternation. During the course of its development, however, IE gradually grammaticalized these ablaut alternations.

IE had constraints on permissible root structure sequences. In terms of the radical revision of the IE consonant system proposed by Hopper, Gamkrelidze, and Ivanov, these constraint laws may be stated as follows:

1. Each root contained at least one non-glottalic consonant.
2. When both obstruents were non-glottalic, they had to agree in voicing.

The IE root structure constraint laws thus become merely a voicing agreement rule with the corollary that two glottalics cannot co-occur in a root. Comparison of IE with the other Nostratic branches indicates, however, that the forbidden root types must have once existed. Two rules may be formulated to account for the elimination of the forbidden types:

1. A rule of progressive voicing assimilation may be set up to account for the elimination of roots whose consonantal elements originally did not agree in voicing: _*T_ ~ _*B_ > _*T_ ~ _*P_, _*B_ ~ _*T_ > _*B_ ~ _*D_, etc.
2. A rule of regressive deglottalization may be set up to account for the elimination of roots containing two glottalics: _*T'_ ~ _*K'_ > _*T_ ~ _*K'_, etc. This rule finds a close parallel in Geers' Law of Akkadian.

According to Gamkrelidze, Bartholomae's Law is a later manifestation of the progressive voicing assimilation rule, applied to contact sequences.

C. Verb morphology. Comparison of IE with Uralic reveals many striking similarities in verb morphology and allows us to ascertain the ultimate origin of the athematic verb endings: they can be nothing else but agglutinated personal pronouns. The earliest forms of the athematic endings were most likely as follows:

Person	Singular	Plural
1	_*-m_	_*-me_
2	_*-t_	_*-te_
3	_*-s_, _*-∅_	_*-se_

The 2nd singular ending _*-t_ is preserved in Hittite. This was later replaced by what had been the 3rd singular, namely, _*-s_. Calvert Watkins (1962) discusses the extensive evidence from the IE daughter languages for an original 3rd singular ending in _*-s_. It was Watkins who also showed that the 3rd singular indicative was originally characterized by the fundamental ending _zero_. The 3rd plural was replaced, at a very early date, by a form of nominal origin along the lines suggested by Oswald Szemerényi. The development of the 3rd singular ending *-*t* was a later change, though this still occurred fairly early since it is found in Hittite and the other Anatolian daughter languages. The most recent change must have been the development of the so-called 'primary' endings, which were built upon the so-called 'secondary' endings by the addition of the deictic particle *-*i* meaning 'here and now', as shown by Kerns and Schwartz in their book on Indo-European verb morphology (1972). It may be mentioned that this deictic particle has a Nostratic origin, coming from a widely-represented proximate demonstrative stem meaning 'this one here'.

The comparison with Uralic also shows that the earliest IE had two conjugational types: (A) a determinative (objective) conjugation, which was characterized by the 3rd singular in *-*s* and which was used with transitive verbs, and (B) an indeterminative (subjective) conjugation, which was characterized by the 3rd singular in *zero* and which was used with intransitive verbs. This is identical to what is assumed to have existed in the Uralic parent language.

D. Vowel gradation. The development of vowel gradation is extremely complicated and would require far more space to discuss than is alloted for this paper. Therefore, I will only deal with several key points. Ever since Hirt, it has been assumed by many scholars that early IE went through a stage of development characterized by phonemic stress and that this stress caused the weakening and/or loss of the vowels of unaccented syllables, that is to say that the stress was responsible for the development of the quantitative ablaut alternations. Furthermore, according to this theory, it is assumed that, at a later date, stress became phonemically non-distinctive and was replaced by an accent system characterized by phonemic pitch, and that this pitch accent was responsbile for the development of the qualitative ablaut alternations. Kuryłowicz, however, argued that the qualitative ablaut alternations were ancient and preceded the changes brought about by the phonemicization of a strong stress accent. Comparison with the other Nostratic languages, especially Kartvelian, indicates that Kuryłowizc was correct. IE inherited the qualitative ablaut alternations from Nostratic. Greenberg (1990) supplies convincing evidence in support of this view. The phonemization of a strong stress accent in early IE brought about a complete restructuring of the inherited vowel system. The same thing happened in Kartvelian. Another important point concerns the early prehistory of the **e* ~ **o* ablaut gradation. Pulleyblank (1965) tried to show that this gradation series should be reinterpreted as a **ə* (*schwa*) ~ **a* gradation. It looks as though

Pulleyblank came pretty close to the truth, though only for the oldest period of development. We may note that this older system is partially preserved in Hittite, where **ə* appears as *e* (or *i*) and **a* is preserved as such. The development of **ə* to **e* is fairly easy to explain: **e* may be assumed to have been the normal allophone of **ə* under stress. A typological parallel may be observed in the Northwest Caucasian languages Ubykh and Circassian, where *ə* becomes *e* under stress. For the latest period of development, namely, the period directly before the emergence of the non-Anatolian daughter languages, the traditional system of five long and short vowels is surely correct. Finally, there is little indication that Nostratic had phonemic long vowels. Therefore, long vowels may be assumed to have arisen soley in IE proper.

E. Indo-European consonantism. There are internal inconsistencies in the traditional reconstruction of the IE stop system that make that system highly improbable from a typological point of view. In order to address these problems, Thomas Gamkrelidze and Vjačeslav Ivanov, on the one hand, and Paul Hopper, on the other, independently proposed in 1972 and 1973, respectively, a radical reinterpretation of the IE stop system. According to Hopper, Gamkrelidze, and Ivanov, the traditional plain voiced stops are to be reinterpreted as glottalized stops (that is, ejectives). Furthermore, according to the version of the theory proposed by Gamkrelidze and Ivanov, the traditional plain voiceless stops are to be reinterpreted as voiceless aspirates, while the traditional voiced aspirates are to remain unchanged. In this revised interpretation, aspiration is viewed as a phonemically redundant feature, and the phonemes in question could also be realized as allophonic variants without aspiration. Strong support for this theory is provided by comparison of IE with Kartvelian and Afroasiatic, both of which have a three-way contrast, in the series of stops and affricates, of voiceless (aspirated) ~ glottalized ~ voiced. According to my views on Nostratic, though not according to the views of Illič-Svityč and Dolgopolsky, the IE glottalized stops (the traditional plain voiced stops) correspond exactly to glottalized stops in Kartvelian and Afroasiatic, while the voiceless (aspirated) stops in IE correspond to identical sounds in Kartvelian and Afroasiatic, and the voiced (aspirated) stops of IE correspond to voiced stops in Kartvelian and Afroasiatic. It should be noted that the voiced aspirates were probably a late development in IE, and this series was probably originally characterized by plain voicing, without aspiration.

This paper is merely an overview of current research—we have barely scratched the surface. Nostratic studies are still in their infancy, and there remain many issues to be resolved, but the future looks extremely exciting and extremely promising.

73 Phillips Street
Boston MA 02114

REFERENCES

Birnbaum, Henrik. 1977. *Linguistic reconstruction: Its potentials and limitations in new perspective.* Washington, DC: Journal of Indo-European Studies.

Bomhard, Allan R. 1977. ''The Indo-European/Semitic hypothesis re-examined''. *JIES* 5/1: 55–99.

Bomhard, Allan R. 1981. ''Indo-European and Afroasiatic: New evidence for the connection''. *Bono homini donum: Essays in historical linguistics in memory of J. Alexander Kerns,* part I. Ed. Yoël L. Arbeitman and Allan R. Bomhard. Amsterdam: John Benjamins. Pp. 351–474.

Bomhard, Allan R. 1984. *Toward Proto-Nostratic: A new approach to the comparison of Proto-Indo-European and Proto-Afroasiatic.* Amsterdam: John Benjamins.

Bomhard, Allan R. 1986a. Review of Vitalij V. Shevoroshkin and Thomas L. Markey, eds., *Typology, relationship, and time. Diachronica* 3(2):269–81. (A revised version of this review appeared in *Mother Tongue,* issue 10, April 1990).

Bomhard, Allan R. 1986b. ''Common Indo-European/Afroasiatic roots: Supplement I''. *General linguistics* 26(4):225–57.

Bomhard, Allan R. 1988. ''The prehistoric development of the athematic verbal endings in Proto-Indo-European''. *A linguistic happening in memory of Ben Schwartz: Studies in Anatolian, Italic, and other Indo-European languages.* Ed. Yoël L. Arbeitman. Louvain: Peeters. Pp. 475–88.

Bomhard, Allan R. 1990. ''A survey of the comparative phonology of the so-called 'Nostratic' languages''. *Linguistic change and reconstruction methodology.* Ed. Philip Baldi. The Hague: Mouton. Pp. 331–58. (A Russian language version of this paper also appeared in *Voprosy jazykoznanija* 1988(5):50–65).

Bomhard, Allan R. 1991. ''Lexical parallels between Proto-Indo-European and other languages''. *Studia etymologica Indoeuropaea: Memoriae A. J. Van Windekens (1915–1989) dicata.* Ed. L. Isebaert. Leuven: Peeters.

Bomhard, Allan R. Forthcoming. ''The root in Indo-European and Afroasiatic''. *Gedenkschrift for Karel Petráček.* Ed. Petr Vavroušek and Petr Zemánek.

Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov. 1972. ''Lingvističeskaja tipologija i rekonstrukcija sistemy indoevropejskix smyčnyx''. *Konferencija po sravnitel'no-istoričeskoj grammatike indoevropejskix jazykov, pradvaritel'nye materialy.* Ed. C. B. Bernstein, V. V. Ivanov, V. A. Dybo, and R. V. Bulatova. Moscow: Nauka. Pp. 15–18.

Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov. 1973. ''Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse''. *Phonetica* 27:150–56.

Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov. 1984. *Indoevropejskij jazyk i indoevropejcy: Rekonstrukcija i istoriko-tipologičeskij analiz prajazyka i protokul'tury.* 2 vols. Tbilisi: Publishing House of the Tbilisi State University.

Greenberg, Joseph H. 1957. *Essays in linguistics.* Chicago: Unviersity of Chicago Press.

Greenberg, Joseph H. 1990. ''The prehistory of the Indo-European vowel system in comparative and typological perspective''. *Proto-languages and proto-cultures.* Ed. Vitaly Shevoroshkin, pp. 77–136. Bochum: Universitätsverlag Brockmeyer.

Greenberg, Joseph H. Forthcoming. *Indo-European and its closest relatives: The Eurasiatic family.* Stanford, CA: Stanford University Press.

Hopper, Paul J. 1973. ''Glottalized and murmured occlusives in Indo-European''. *Glossa* 7: 141–66.

Illič-Svityč, V(ladislav) M(arkovič). 1965. ''Materialy k sravnitel'nomu slovaru nostratičeskix jazykov''. *Etimologija* 1965:321–73.

Illič-Svityč, V(ladislav) M(arkovič). 1971-. *Opyt sravnenija nostratičeskix jazykov (semitoxamitskij, kartvel'skij, indoevropejskij, ural'skij, dravidskij, altajskij).* 3 vols. Moscow: Nauka.

Kerns, J. Alexander and Benjamin Schwartz. 1972. *A sketch of the Indo-European finite verb.* Leiden: E.J. Brill.

Pedersen, Holger. 1931. *The discovery of language: Linguistic science in the nineteenth century.* Tr. John Webster Spargo. Midland Book, 1962. Bloomington: Indiana University Press.

Pulleyblank, Edwin G. 1965. ''The IE vowel system and qualitative ablaut,'' *Word,* 21:86–101.

Watkins, Calvert. 1962. *Indo-European origins of the Celtic verb I. The sigmatic aorist.* Reprinted 1970. Dublin: Dublin Institute for Advanced Studies.

BERNARD H. BICHAKJIAN

Genealogical linguistics: Progress and pitfalls

Abstract. The classification of unwritten languages, evidence from archaeology, and parallel research in population genetics have revived the interest in genealogical linguistics and expanded its methods. The development can only be welcomed, provided, of course, the potential of each method is properly recognized. The use of the lexical-resemblance criterion for uncovering patterns of kinship among languages provides a possibility for exploring areas which the comparative method could not reach. But patterns of kinship are one thing, ancestral forms another. This paper will argue that, without recognizing the principle of a developmental process in linguistics and without having access to the corresponding empirical data, reconstruction efforts are doomed to produce insufficient and misleading results.

1. A necessary compass. In a methodologically important article, published in 1957 and endorsed by Jakobson, who reemphasized its central idea at the CIPL conference meeting that year, Greenberg argued that the reconstructed shape of a protolanguage cannot be plausible unless it conforms in the first place to the principles that regulate the organization of extant linguistic models. This basic acceptability condition led to growing misgivings about Brugmann's reconstruction of the IE stops and paved the way to Gamkrelidze and Ivanov's glottalic theory (1973).

Today, as interest in formal analyses wanes and corroborating evidence from archaeology and molecular biology provides new incentive, linguists are seeking to lift the veil that shrouds the ancestry of protolanguages. As a result, the ultimate genealogical tree, representing the derivational pattern of all human languages from a common ancestral tongue, has been proposed, and attempts are made to reconstruct isolated items of the original lexicon (for an excellent presentation of the state of the art, see Wright, 1991). The willingness to address unsolved problems is of course welcome and, given the lack of observable linguistic data from the period that is dealt with, the indications provided by bordering disciplines, especially population genetics, are indeed a precious asset. Manaster-Ramer is quite right to point out that "genetics and linguistics should jointly yield a consistent picture of prehistory . . . [and] the idea that there may have been a language . . . from which all the surviving languages on this planet are derived, must be derived from non-linguistic evidence" (1989:106–7). Yet, the reborn pioneering spirit and the collateral effort it has enlisted run the risk of not fulfilling the promise they hold because of a serious operational oversight.

To understand the problem, one must recognize that we are dealing with two distinct periods. The first one stretches from ca. 92,000 BP, when the initial division is estimated to have taken place in the population of modern humans (Cavalli-Sforza et al. 1988:6004; for a discussion of the congruence between demographic and linguistic genealogical trees, see Bateman et al. 1990, and for additional evidence Excoffier et al. 1991), and, therefore, presumably when Protoworld lost its unity, to the dialectalization of the traditional protolanguages (cf. e.g. Ruhlen 1987:275 et sq.), which, going by IE and using Gamkrelidze and Ivanov's chronology, could be set around 6000 or 7000 BP (1985:26; also cf. Polomé 1982). The second period covers the millennia that separate the first one from the present. Though reconstructed forms are undeniably putative, the reliability of the comparative method and the abundance of the material, especially in the well documented language families, suggest that it is not unreasonable to put the recorded and reconstructed data together and consider the second period the observational period and the first one the conjectural one.

The above-mentioned operational oversight can now be understood. Prospecting the conjectural period is certainly a legitimate pursuit, but can it be done without first assessing what transpired during the observational period? Would the chances of error not be markedly reduced by making surmises constrained by the evidence provided by the empirical data? These rhetorical questions bring us back to Greenberg's condition. Just as reconstructed models must comply with the organizational principles of existing languages, so must the surveying of the conjectural period be placed in the developmental perspective that manifests itself in the observational period. This is precisely where the oversight lies. Linguists have insufficiently recognized the linear process that has unfolded during the recorded and reconstructed history of languages, and by so doing have deprived themselves of a valuable compass for charting their research into the remote phases of the conjectural period.

2. Developmental approaches in linguistics. The study of the developments that have taken place during the observational period has had a checkered history. Schleicher's ill-conceived, but explainable, assumption of linguistic decay ('Verfall') is well-known (cf. Schleicher 1852:14–30; and for a discussion Leroy 1950:10–20), and so is Jespersen's better-inspired, but beleaguered theory of linguistic progress (1964:364; for the problems surrounding the notion of progress in biology, see Nitecki 1988 and therein Ruse's candid discussion), with in between the descriptive accounts of Meillet, who saw converging processes (1965), and Sapir, who observed a limited number of major drifts (1949:147 et sq.). While interpretations and metaphors vary, the facts remain. Many of them were presented at another CIPL conference by Giuliano Bonfante, who remarked that "le temps est venu, je crois, de commencer à esquisser une méthode qui nous permettrait d'établir les lois du développement progressif du langage humain" (1939:31). The issue may have been ripe, but the times were less than propitious—the paper had been prepared for a plenary session scheduled for the week of August 28 through September 2, 1939! When the guns fell silent, concerns had shifted, and Bonfante's call was left unheeded. Today, a few wars later and at the dawn of a new world order, it may not be inappropriate to reexamine the issue of "le développement progressif du langage humain."

2.1 Choosing the right approach. Strangely enough, the empirical data are not in short supply. The IE family of languages, which offers the broadest and best explored field of observation, provided Bonfante with no less than twelve unidirectional changes, including the loss of the dual, the creation [sic] of articles, the shift from aspect to tense, and the reversal of the *déterminant-déterminé* word order. But the enumeration of apparently disparate data does not prove anything. It is instead a self-defeating exercise, because it allows opponents to refute the original claim by alleging seemingly contradictory data, which can always be found. This is precisely what happened when Collinder argued that the history of certain Hungarian morphological items has run contrary to the shift toward greater analyticity generally associated with the IE and Hamito-Semitic languages (1936:58–59 and 1956:120). The dismantling of declensions was no longer a foreordained finality—one day the IE affixes would be reinstated in their venerable glory, just as, in Lyell's vision, "extinct genera would reappear when the appropriate stage of the grand climatic cycle, or great year, came round again" (Gould 1988:323).

Though many hastened to conclude that the Hungarian data delivered a fatal blow to the developmental argument (cf. e.g. Malmberg 1968:26), in true fact Collinder's observations only revealed the insufficiency of the enumerative method. A discussion exclusively conducted at the observational level can only degenerate into a sterile exchange of examples and counterexamples; if they are to fulfill their heuristic promises, the competing hypotheses must be directed instead at the underlying process and the guiding principle that account for the recorded occurrences. It will be recalled that the theory of evolution is not primarily associated with Lamarck, the author of the first comprehensive treatise on transmutation, or with Mendel, the pioneer in heredity research, but with Darwin, who understood that the notion of biological evolution could not be accepted unless the operating principle is uncovered and defined; hence the importance he gave to natural selection in his theory and in the title of his treatise, *On the origins of species* BY MEANS OF *natural selection* . . . (emphasis added).

It is not enough therefore to draw a long list of seemingly unidirectional changes to prove that languages develop. Such an inventory is only the first step. The essential task is to uncover the general process to which these individual changes can be reduced and find the operating principle that governs it. The history of the future from early Latin to present-day French will provide a telling illustration.

2.2 The evidence from the Latin and French features. If the discussion is limited to a superficial comparison of forms such as *cantābō, cantāre habeō, (je) chanterai* and *(je) vais chanter,* one will promptly conclude that the history of the Latin and French futures is the alternating occurrence of analytic and synthetic structures. This is so, in part at least, and one should not try to hide it. But there is more to the history of the future in these languages.

1. The original structure of *cantā-bō* was indeed analytic, but one should not oversee the significance of the constituents. The first one was the imperfective form of the verb, while the other was the present of the verb 'to be.' These two features suggest that at least when it was coined, this form had an aspectual value (cf. its exact equivalent in the Russ. *ja budu pet'*).

2. The originality of *cantāre habeō* now comes to light. It is made of an aspectually-neutral infinitive—gone is the imperfectum/perfectum opposition—and of 'to have,' the auxiliary of temporal conjugations. The Late Latin future was not therefore simply an analytic

form—it was a tense, whereas its ancestor had been part of an aspectual system.

3. The increasingly pervasive (*je*) *vais chanter* is also more than just an analytic alternative. Whereas *cantāre habeō* was a head-last or SOV-type structure, (*je*) *vais chanter* is a head-first periphrasis of the SVO type.

4. Returning to Latin, one may finally point out that *canet,* the future of the original form, was derived from an ancestral subjunctive.

The history of the future cannot therefore be reduced to a hide-and-seek game between analytic phrases and synthetic forms, and only superficially does it provide a linguistic example of Hutton's vision of an endless sequence of cycles that affords "no vestige of a beginning, no prospect of an end." The essential changes are elsewhere: The ancestral modo-aspectual distinctions have given way to a temporal one, and the head-last structures, also characteristic of older linguistic stages, are shifting to their symmetrical equivalents (for a less sketchy account of the history of the future, see Bichakjian 1990; and, for a discussion of the IE verbal system, Bichakjian 1988b).

Though the history of the future has now been reduced to two major changes—the transition from aspect to tense and the shift from head-last to head-first structures, which are widely spread among the languages of the world (cf. Givón 1979 and Bichakjian 1991), and which *mutatis mutandis* already appeared on Bonfante's list—the two phenomena seem to be distinct. Further examination will reveal that they in fact share an important property.

2.3 The underlying process. When the acquisition of temporal distinctions by children learning their native language is compared to the mastery of aspectual oppositions by similar learners in similar situations, one realizes that temporal distinctions are acquired earlier than aspectual ones (cf. e.g. Stéphany 1981:185). A similar comparison of head-first and head-last structures also reveals that the acquisition of one lags behind that of the other. By and large and especially when they become longer, head-first structures are definitely acquired earlier than their symmetrical alternatives (cf. e.g. Sheldon 1972:9–10 and Bichakjian 1987).

The chronology of the acquisition of linguistic features now emerges as an important indicator, and, when the major changes are surveyed, it appears consistently that the outgoing features were acquired late in childhood, while the alternatives that came in their stead are mastered at a younger age. In the IE languages, the late-acquired ancestral laryngeals and the preceding *e* were replaced with long vowels, that are acquired earlier, and, when these long vowels in turn were eliminated, the systems were sometimes expanded with front rounded or back unrounded vowels, that appear even before quantitative distinctions are made systematically. In addition to the shift from aspect to tense and from head-last to head-first structures, the IE languages have developed subordinate clauses, which are acquired earlier than their participial forerunners, and a variety of free morphemes such as prepositions, pronouns, articles, and degree adverbs, whose mastery takes place before the complete sets of their ancestral alternatives (for a less cursory account, see Bichakjian 1988a). The IE languages have evolved by replacing late-acquired features by ever-earlier acquired alternatives.

2.4 The operating principle. The chronological evidence from language acquisition now makes it possible to reduce all the major changes to a single well-defined process. But although the shift to earlier acquired features does capture the unity of the changes observed in phonology as well as morphology and syntax, the process may seem arbitrary unless the guiding principle is uncovered. When a linguistic feature is replaced by an earlier acquired alternative, the speaker is provided with a number of significant advantages. The corresponding neural pathways are presumably simpler and demonstrably stronger (cf. *mutatis mutandis* LeVay, Wiesel, and Hubel 1980), which makes them more economical to operate, more impervious to somatic or psychological strain, and less vulnerable to accidental or degenerative aphasia. Moreover earlier-acquired features provide earlier-acquired proficiency in language, which in turn fosters earlier and better-established socialization and earlier and fuller mental development.

The total picture is beginning to emerge. While the psycholinguistic data suggested above that the underlying process is a constant shift to ever-earlier-acquired features, the indications and arguments from developmental neurology now offer a possible explanation. Languages proceed in the direction of earlier-acquired features because such a shift provides important selective advantages.

The question must not be therefore whether a given change is universal—it does not have to be because it may be advantageous to introduce a given feature here and eliminate it there, cf. e.g. vowel length, which is acquired earlier than laryngeals, but later than *y* or *ɨ*—but rather whether languages proceed in one general direction, namely in the direction of earlier-acquired features. Since the opposite process would introduce greater and greater handicaps, the general shift to ever-later acquired features seems a very unlikely alternative. That is why Hungarian and the other Finno-Ugric languages have on the whole replaced late-acquired items with earlier acquired alternatives (Bichakjian, 1991). Of course, languages have not developed at the same pace, and the shift to earlier-acquired features is only a general direction, but as Sapir had keenly observed there is direction (1949:155), and the present use of psycholinguistic data and neurological arguments define and explain that direction.

While the linguistic evidence is clear and the increased preservation of juvenile features is well known in hominid evolution (cf. e.g. Gould 1977) and emphatically recognized in ethology (Lorenz 1971), there are linguists, who are uncomfortable with the notion of development. As a school, Transformational Grammar is in the Uniformitarian frame of mind characteristic of pre-Darwinian England, and many linguists, independently of their theoretical affiliation, peremptorily assert that there is no evolution in language, and that changes are exclusively cyclical. It is fair to ask when these ideas, whose ''fallacious'' character was denounced by Sapir some seven decades ago (1949:155), and whose equivalents in the natural sciences were tartly ridiculed in the nineteenth century (see in fig. 1 De la Beche's caricature of Lyell's cyclical conception of evolution) before being irrevocably rejected, will be abandoned, but the prediction is difficult to make. Biological evolution was not universally accepted overnight, either. The only certainty today is that those who see only cyclical changes in the history of languages have a serious impediment: They cannot separate the wheat from the chaff, a handicap often compounded with overbearing reactions, while modesty would be more fitting.

3. Putting the past in perspective. Having argued that languages have proceeded during the observational period toward ever earlier-acquired features, I may now return to the conjectural period. This is indeed a period for which we can only make conjectures, but the safest one is that the process that was inferred for the observational period must have been at work from the very beginning. We cannot tell how fast it moved during this presumed 85,000-year span—like technological advance, it probably progressed at an accelerated rate, with some branches showing more acceleration than others—but there can be little doubt that the features used by the speakers of Protoworld were acquired commensurately later than those we use today and even later than those we reconstruct for the traditional protolanguages.

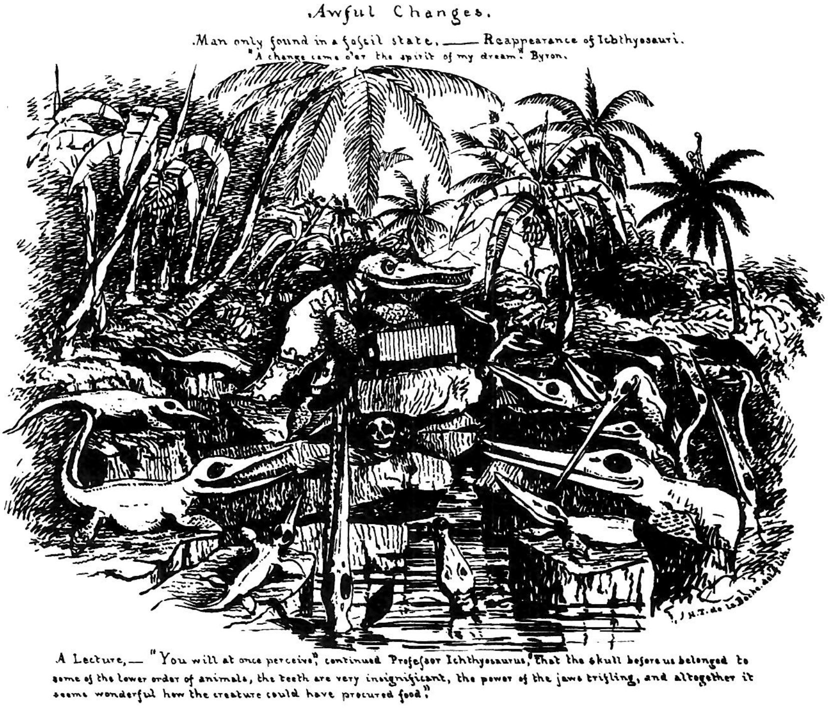

Awful Changes.
Man only found in a fossil state. — Reappearance of Ichthyosauri.
"A change came o'er the spirit of my dream." Byron.
A Lecture, — "You will at once perceive," continued Professor Ichthyosaurus, "that the skull before us belonged to some of the lower order of animals, the teeth are very insignificant, the power of the jaws trifling, and altogether it seems wonderful how the creature could have procured food."

Figure 1. ''Awful Changes.'' ''Man only found in a fossil state,—Reappearance of Ichthyosauri. 'A change came o'er the spirit of my dream.' Byron.'' ''A Lecture,—'You will at once perceive,' continued Professor Ichthyosaurus, 'that the skull before us belonged to some of the lower order of animals, the teeth are very insignificant, the power of the jaws trifling, and altogether it seems wonderful how the creature could have procured food.' '' (Reproduced from Rudwick, 1975:539).

3.1 Recognizing developmental history. It is tempting, of course, to speculate about the very nature of the original features. The fascination of humans for the Adamic language is well-known, and so are the risks of ridicule. Let us again take the example of the Romance future. If the Latin form had to be reconstructed from extant data, it would be absolutely impossible to go beyond the *cantāre habeō* stage—the use of the auxiliary 'to be' with the imperfective verbal noun would never be revealed, and the modal source of the older *canet* would not be suspected. In other words, the developmental history of the future from an ancestral subjunctive to an aspectually marked tense and finally to a strictly temporal distinction would not come to light, and some might be led to assume that the future tense as it is today has always existed as such. This is not a call to inaction or defeatism, it is a call to realize that the linguistic features that exist today or that we reconstruct for the traditional protolanguages have a developmental history behind them, and that we cannot with impunity transpose them as such into a language assumed to have been spoken some 90,000 years ago.

3.2 Learning from past errors. The lack of a developmental perspective was blatantly exposed at the beginning of this century, when British palaeontologists refused to accept that the skullcap of the Javanese Pithecanthropus or the head of the juvenile Australopithecus found at Taung, South Africa, could be the fossil remains of an ancestral species intermediate between modern humans and extant anthropoids, the famous missing link. They made instead the enormous blunder of maintaining for some forty years that the true specimen of the human ancestor was the Piltdown man, who turned out to be a cruel hoax. A clever forger had stained the jawbone of an orangutan, filed down its teeth and deposited it along with fragments of a modern human cranium in a geological site dating back to early Pleistocene. Of course, the carbon-14 method was not available, and national pride played a disproportionate role, but the reason why, as Johanson and Shreeve put it, "so many trained scientists were so utterly and shamefully fooled" is because the Javanese "skullcap was too thick and primitively featured to satisfy the expectations many scientists had for an ancestor of man" (1989:52 and 49), and the Taung child was disqualified for the same reason. Sir Arthur Keith, one of the leading anatomists of his time was categorical: "The skull is . . . showing so many points of affinity with . . . the gorilla and the chimpanzee, that there cannot be a moment's hesitation in placing the fossil form in this living group" (quoted from Johanson and Shreeve 1989:56). In his skillful analysis of the rejection of Taung and the acceptance of Piltdown, Lewin reaches the same conclusion as Johanson and Shreeve: "How is it that trained men, the greatest experts of their day, could look at a set of human bones—the cranial fragments—and 'see' a clear simian signature in them; and 'see' in an ape's jaw the unmistakable signs of humanity? The answers, inevitably, have to do with the scientists' expectations and their effects on the interpretation of data" (1987:61). The expectations and the corollary objective could not have been clearer nor indeed more ludicrous—humans were to be given an ancestor whose essential features would be those of today's population. The gravity of the error becomes apparent when one realizes that evolutionists were ready to deny evolutionary principles in order to preserve their perennial picture of humans.

3.3 The lesson. For linguists, the lesson seems obvious. Uncovering the developmental changes and the underlying process that have been active during the observational period seems to be an important step; combining the linguistic data with the archaeological and genetic indications to outline genealogical patterns during the conjectural period is certainly a legitimate endeavor; but reconstructing a Protoworld whose features would be those of today's languages is a task that the linguist who has learned the Piltdown lesson will wisely eschew.

University of Nijmegen
Erasmusplein 1
6525 GG Nijmegen
The Netherlands

REFERENCES

Bateman, Richard, Ives Goddard, Richard O'Grady, V.A. Funk, Rich Mooi, W. John Kress, and Peter Cannell. 1990. "Speaking of forked tongues. The feasibility of reconciling human phylogeny and the history of language." *Current anthropology* 31:1–24.

Bichakjian, Bernard H. 1987. "The evolution of word order: A paedomorphic explanation." *Papers from the Seventh international conference on historical linguistics*. Eds. Anna Giacalone Ramat, Onofrio Carruba and Giuliano Bernini. Amsterdam: Benjamins. Pp. 87–107.

———. 1988a. *Evolution in language*. Ann Arbor, MI: Karoma.

———. 1988b. "**J'ai tombé* pour *je suis tombé:* L'aboutissement d'une longue évolution." *Aspects de linguistique française. Hommage à Q.I.M. Mok.* Ed. Ronald Landheer. Amsterdam: Rodopi. Pp. 31–48.

———. 1990. "Language change: Cyclical or linear? The case of the Romance future." *Historical linguistics 1987. Papers from the Eighth international conference on historical linguistics*. Eds. Henning Andersen and Konrad Koerner. Amsterdam: Benjamins. Pp. 37–49.

———. 1991. "Evolutionary patterns in linguistics." *Studies in language origins, II.* Eds. Walburga von Raffler-Engel and Jan Wind. Amsterdam: Benjamins. Pp. 187–224.

Bonfante, Giuliano. 1939. "Les lois du développement linguistique." *Réponses au questionnaire (suite) du Cinquième congrès international des linguistes,* Bruxelles, 1939. Bruges: Imprimerie Sainte Catherine. Pp. 31–32.

Cavalli-Sforza, Luigi Luca, Alberto Piazza, Paolo Menozzi, and Joanna Mountain. 1988. "Reconstruction of human evolution: Bringing together genetic, archaeological, and linguistic data." *Proceedings of the National Academy of Sciences USA* 85:6002–06.

Collinder, Björn. 1936. "Analytische Sprachentwicklung und linguistische Teleologie." *Nyelvtudományi Közlemények* 50.51–63.

———. 1956. "Die Entstehung der Sprache." *Ural-Altäische Jahrbücher* 28:116–27.

Excoffier, Laurent, Rosalind M. Harding, Robert R. Sokal, Beatrice Pellegrini, and Alicia Sanchez-Mazas. 1991. "Spatial differentiation of Rhesus and Gm haplotype frequencies in Sub-Saharan Africa and its relation to linguistic affinities." *Human biology* 63:273–307.

Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov. 1973. "Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse." *Phonetica* 27:150–6 [first published in Russian in *Working papers of the Conference on the comparative grammar of the Indo-European languages*. Moscow, 1972].

———. 1985. "The ancient Near East and the Indo-European question: Temporal and territorial characteristics of Proto-Indo-European based on linguistic and historic-cultural data." *Journal of Indo-European studies* 13:3–48.

Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.

Gould, Stephen J. 1977. *Ontogeny and phylogeny*. Cambridge, MA: The Belknap Press of Harvard Univ. Press.

———. 1988. "On replacing the idea of progress with an operational notion of directionality." In Nitecki, ed. Pp. 319–38.

Greenberg, Joseph H. 1957. "The nature and uses of linguistic typologies." *International journal of American linguistics* 23:68–77.

Jakobson, Roman. 1958. "Typological studies and their contribution to historical comparative linguistics." *Proceedings of the Eighth international congress of linguists*, Oslo, 1957. Ed. Eva Sivertsen. Oslo: Univ. Press. Pp. 17–35 [also in *Selected writings*. The Hague: Mouton, 1962. I, 523–32].

Jespersen, Otto. 1964. *Language: Its nature, development, and origin*. New York: Norton [first publ. in 1922. London: Allen and Unwin; New York: Holt].

Johanson, Donald, and James Shreeve. 1989. *Lucy's child: The discovery of a human ancestor*. New York: Avon.

Leroy, Maurice. 1950. *Sur le concept d'évolution en linguistique*. Bruxelles: Office de Publicité [originally publ. in *Revue de l'Institut de Sociologie*, 3 (1949):337–76].

LeVay, Simon, Torsten N. Wiesel, and David H. Hubel. 1980. "The development of ocular dominance columns in normal and visually deprived monkeys." *The journal of comp. neurology* 191:1–51.

Lewin, Roger. 1987. *Bones of contention: Controversies in the search for human origins*. New York: Simon and Schuster.

Lorenz, Konrad. 1971. *Studies in animal and human behaviour*. Vol. 2. Tr. Robert Martin. London: Methuen.

Malmberg, Bertil. 1968. *Les nouvelles tendances de la linguistique*. Tr. Jacques Gengoux. 2nd ed. Paris: PUF [original: *Nya vägar inom språkforskningen. En orientering i modern lingvistik*. Stockholm: Svenska Bokförlaget, 1962].

Manaster-Ramer, Alexis. 1989. "Language, history and computation (preliminary report)." *Reconstructing languages and cultures*. Ed. Vitaly Shevoroshkin. Bochum: Brockmeyer.

Meillet, Antoine. 1964. *Introduction à l'étude comparative des langues indo-européennes*. University, AL: Univ. of Alabama Press [a rpt. of the 8th ed. Paris: Hachette, 1937].

———. 1965. "Convergences des développements linguistiques." *Linguistique historique et linguistique générale*. 2nd ed. Paris: Champion. Pp. 61–75 [first publ. in *Revue Philosophique* 85 (1918)].

Nitecki, Matthew H., ed. 1988. *Evolutionary progress?* Chicago: The Univ. Press.

Polomé, Edgar C. 1982. *The Indo-Europeans in the fourth and third millennia*. Ann Arbor, MI: Karoma.

Rudwick, Martin J.S. 1975. "Caricature as a source for the history of science: De la Beche's anti-Lyellian sketches of 1831." *Isis* 66:534–60.

Ruhlen, Merritt. 1987. *A guide to the world's languages.* Vol. 1. *Classification.* Stanford, CA: UP.

Ruse, Michael. 1988. "Molecules to men: Evolutionary biology and thoughts on progress." In Nitecki, ed. Pp. 97–126.

Sapir, Edward. 1949. *Language: An introduction to the study of speech.* New York: Harcourt, Brace and World [first published in 1921].

Schleicher, August. 1852. *Les langues de l'Europe moderne.* Tr. Hermann Ewerbeck. Paris: Ladrange and Garnier [German original: *Die Sprachen Europas in systematischer Übersicht.* Bonn: König, 1850; also published as vol. II of *Sprachvergleichende Untersuchungen.* Bonn: König, 1848–50].

Sheldon, Amy Louise. 1972. "The acquisition of relative clauses in English." Doctoral Diss. The Univ. of Texas at Austin.

Stephany, Ursula. 1981. *Aspekt, Tempus und Modalität: Eine Studie der Entwicklung der Verbalgrammatik in der neugriechischen Kindersprache.* 2 vols. Preprint version [Stephany's research is reported in part in "Modality." *Language acquisition: Studies in first language development.* 2nd ed. Eds. Paul Fletcher and Michael Garman. Cambridge, The Univ. Press, 1986. Pp. 375–400].

Wright, Robert. 1991. "Quest for the mother tongue." *The Atlantic monthly* 267 No 4 (April): 39–68.

Notes and Discussion

On the etymology of Portuguese *SAUDADE:* an instance of multiple causation?

LEO PAP

One of the most discussed words in the Portuguese language is the noun *saudade.* It refers to a kind of emotion often said to be distinctively rooted in the Portuguese national character, so much so that it has been declared untranslatable; an approximate gloss would be 'nostalgic longing'. The following discussion is more concerned with the formal derivation than the exact semantic development of *saudade.* Beyond dealing with a term of particular cultural interest to Lusophiles, this article touches on the more general question of "multiple causation" in establishing etymologies, applicable to many other words in Portuguese as well as in other languages.

The two leading etymological dictionaries of the Portuguese language, by Antenor Nascentes and by Francisco da Silveira Bueno, derive *saudade* from Lat. *solitate(m)* 'loneliness, forlornness', which yielded the archaic Port. forms *soedade, soidade, suidade* (or *soydade, suydade*); these archaic forms then came under the influence of Port. *saudar* (<Lat. *salutare* 'to wish good health, to hail, to greet').[1] Authorities cited by Nascentes include such notables as Friedrich Diez, Wilhelm Meyer-Lübke, Adolpho Coelho, José Leite de Vasconcelos, Aniceto Gonçalves Viana, among others. Moreover, Carolina Michaëlis de Vasconcelos has devoted an entire monograph to a discussion of the Portuguese notion of *saudade,* including the etymology of the word.[2]

Michaëlis distinguishes six stages in the transition from Lat. *solitate(m)* to modern Port. *saùdade:* (1) *solitate,* (2) *soedade,* (3) *soïdade,* (4) *suïdade* (or *ssuydade*), (5) *saùdade,* (6) *sòdade.* The first four stages involve regular phonetic changes, but the fifth can only be explained, according to Michaëlis, by an "association of ideas and popular etymology". The sixth stage, mentioned in passing, is a popular pronunciation characteristic of Lisbon (representing regular phonetic change: *au* > *ou,* and also, even though Michaëlis does not say so, semantic influence of the adjective *só* 'alone'). The form *soidade,* according to Michaëlis, is attested for the 13th century. The earlier *soedade* survived in Galicia until the 15th century; *suidade* is common today in popular Gallego. In Portugal proper, the form *suidade* was preferred by classical writers until c. 1580.

The noun form *saudade* (and the adj. *saudoso,* haplologically derived from *saudade*) gradually acquired currency in the late 16th century, Michaëlis believes; it spread from the mouths of *semi-cultos* ('half-educated') to the upper classes, but did not gain general acceptance until the early 17th while Portugal was under the rule of Spain. (In Camoẽs we still find *soidade* side by side with *saudade,* without difference in meaning.)

Michaëlis rejects the notion that an awareness of such learned vs. popular doublet forms as *causa/coisa* or, along a historical axis, *auru* (cf. *áureo)/oiro* had anything to do with the change from *soidade* to *saudade.* But she does suggest, p.62, that the substitution of *au* for *oi* provided a more melancholy sound for a concept of sadness! The decisive factor, she believes, was the semantic association with words beginning with *saud-*, viz., *saudar* 'to greet, to wish well', *saudações* 'greetings', *saúde* 'health', and the learnèd Biblical *saudade* (< Lat. *salutate*) 'salvation'. But how, Michaëlis asks, do you get from a state of health and haleness to lonely melancholy? Her answer: In letters and thoughts, one often wishes good health to distant loved ones; a devout religious person may even wish salvation (a good even greater than health) for someone sorely missed—especially the kind of salvation that involves reunion. Thus, according to Michaëlis (pp.66–67), the modern *saudade* has connotations of *saude* wished to those absent, of *saudações* exchanged with them, of a feeling of loneliness due to their absence, and of a desire of ultimate reunion through salvation. For good measure Michaëlis adds a possible semantic association with a pre-modern *saidade* (< Lat. *sanitate,* 'health', cf. modern Port. *são* 'healthy'); thus *soidade* > *saidade* > *saudade*—which I myself consider most unlikely.

On the other hand Cláudio Basto (in *Revista lusitana* 17: 275, 18: 178) attributes the transformation of *soidade* to *saudade* to literary influence: the change from *oi* to *au* may have been due, in part at least, to a "false Latinization", i.e. an overcorrection by men of letters on the assumption that the *oi* in *soidade* represented a popular mispronunciation of *ou* (cf. *oiro*<*ouro*), and that the *ou*, in Latinized literary style, should be reduced to the original *au* (cf. *ouro* 'gold'<Lat. *auru,* with the latter surviving in Port. *áureo* 'golden').

Still another etymological suggestion, cited by Nascentes, is the one by João Ribeiro: *saudade* might derive from Arabic *saudá,* which in Classical Arabic means 'black bile', 'hypochondria', 'melancholy', related to the adjective *aswad* 'black'. (The term *melancolia* 'melancholy' itself, of course, derives from the Greek for 'black bile'.) In Ribeiro's words, Ar. *saudá* literally refers to the 'blackened' or 'bruised' blood within the heart, and, figuratively, to a feeling of profound sadness. In (Arabic) medicine, Ribeiro adds, *saudá* is a liver ailment which has as its symptom a bitter and melancholy sadness. One might object, Ribeiro acknowledges, that Arabic loanwords in Portuguese rarely have to do with feelings. However, he points out, *saudá* refers not only to a feeling, but to a disease, and there are "many" Portuguese names for illnesses actually borrowed from Arabic (medieval medicine was dominated by Arabs!); such as *achaque* 'chronic disease, distemper, vice', *enxaqueca* 'migraine headache, hemicrania', and—*soda* 'headache'. (Strangely enough, Ribeiro as cited by Nascentes mentions *soda* without apparently realizing that *soda* in the sense just mentioned does almost certainly derive from Ar. *saudá*!) The exact formal derivation of *saudade* from *saudá* is not indicated in this etymology. Since nouns ending in stressed *-á* are foreign to standard Portuguese one would have to assume a more or less "popular" reinterpretation of *saudá* under the influence of the common feminine noun type containing the *-dade* suffix, and specifically under the influence of Latin-derived *saudade* in its earlier meaning of 'salvation'.

So here we have three different etymologies for *saudade* (in its modern sense): (1) a semantic association of thoughts of greeting, well-wishing, and even of "salvation", with feelings of loneliness and longing; (2) learnèd literary overcorrection based on an awareness of certain popular phonetic changes; (3) morphemic adaptation of a loanword from Arabic, also involving a semantic association with certain similar-sounding native words. Which of these etymologies is correct? Are they mutually exclusive?

In 1956, in a review article of mine on the local speech of Rio de Onor on the Spanish-Portuguese border, I had remarked: ''In etymologizing, one should never overlook the possibility that a given change was produced, not by one factor, but by several converging and mutually reinforcing influences . . .'' (Pap 1956:741 note 8).[3] This issue, of the possible *multiple* causation, i.e. of several contributing and mutually reinforcing factors, rather than just one single cause of some lingual[4] change, was taken up in forceful detail some ten years later by Yakov Malkiel, with reference to the phonological derivation of Spanish adjectives ending in unstressed *-io* (e.g. *limpio* < Lat. *limpidu,* but *amplio,* rather than *amplo* < Lat. *amplu*). (Malkiel 1967a:235–43, and 1967b).[5] Again, in 1978, the question of multiple causation in etymology was discussed, under the label of ''lexical polygenesis'', by Roger W. Wescott, in a paper on such English words as *to slurp, bub, Al.* In Wescott's words, ''most of our overt language behavior is accompanied by a flurry of covert word-association activity'', and where enough people engage in the same kind of multiple word-association over a period of time, ''this eventually reshapes our lexicon.'' Thus, according to Westcott, many words may have more than one lexical ancestry.[6] (Westcott 1978) Finally, in an article on the etymology of Eng. *tip* 'gratuity', published in 1982, I reviewed four different proposed derivations of this word in the indicated sense. I was driven to the conclusion that two or more of these probably reinforced each other; at the same time I conceded: ''Etymologizing, like any historical reconstruction, can only establish probabilities.'' (Pap 1982)[6]

In the case of Port. *saudade* 'nostalgic longing', how did the vowel sequence or diphthong underlying the spellings *oe/oi/oy/ui/uy,* in the immediate derivatives from Lat. *solitate,* turn subsequently into *au?* I believe, although I shall probably never be able to prove it, that here we have another prime example of multiple causation, along the following lines: Probably the association, on semantic grounds, of *soedade* (< Lat. *solitate*) with the several words deriving from Lat. *salut-*, as suggested by Michaëlis, did play a major role, and may even have been the decisive factor with many, but not necessarily with all, speakers of Portuguese around the turn from the 16th to the 17th century, beset by painful historic events. However, the phonemic change from /oy/ to /aw/ seems too ''unnatural'', in the Portuguese context, to have occurred in a single step without considerable resistance. Yet it may very well have been helped along, among the small but influential minority of upper-class literati, who were steeped in Latin, by an awareness of such formal equations as *causa>cousa>coisa, aur-* (as in *áureo,* a literary term)>*ouro*>*oiro, laur-* (as in *láurea*)>*louro*>*loiro.* This would have encouraged the "overcorrective" step from *soidade* to *soudade,* at least; and for those educated folk disapprovingly aware of the popular Lisbonese pronunciation *sòdade* for *saudade* "salvation" etc., it would have been a small further step from *soudade* 'loneliness' to *saudade.* Subconsciously, perhaps, some poetically inclined Portuguese may also have been attracted by the sound symbolism of *au*/aw/, suitable for the expression of "dark" sorrow. A subconscious mental association of *saúd-* and derivates ('hale, hail!, salvation') with Arabic *saudá* 'melancholy' may have provided a reinforcing motivation, for those Portuguese still familiar with old medical terminology, toward altering the meaning of *saudade* 'hale, salvation' in the direction of melancholy longing, somber sadness.

In other words, and in summary, the path of transition from *soedade* (*soidade,* etc.) to *saudade* 'loneliness, melancholy longing', in terms of mental associations, probably was not exactly the same for all Portuguese speakers, during those decisive decades of change that roughly coincided with Spain's rule of Portugal. The etymology proposed by Michaëlis de Vasconcelos may be the "correct" one with reference to a majority of Portuguese speakers; but the etymologies posited by Basto and by Ribeiro may have been, if not decisive, at least contributing factors, with respect to relatively "educated" segments of Portugal's population.

SUNY College at New Paltz
New Paltz, NY 12561

ENDNOTES

[1] Nascentes, in his abridged one-volume edition only, through some oversight or for some deliberate reason, cites the influence of *saúde* 'health' rather than of *saudar.* In the underlying Latin, of course, *salutare* 'to hail' was a verb derived from the noun *salute(m)* 'health'.

[2] The fifth and sixth chapters, pp.55–76, are more specifically devoted to the etymology of the word, *saudade.*

[3] Specifically, Dias and Herculano de Carvalho had noted a recent change from diphthong to simple vowel in certain words, which they attributed to spontaneous internal development. At the same time they had found an inverse shift from [e] to [i̭e] in other words, attributing this diphthongization in some cases to the influence of a preceding word ending in [i], in other cases to a desire to avoid hiatus, and in still other cases to "overcorrection" under standard-Spanish influence. My own conclusion at the time was that Portuguese influence may have *reinforced* an independent tendency among some Rio de Onor dialect speakers to monophthongize; and the diphthongal forms may have been the cumulative result both of overcorrection and of a desire to avoid hiatus, etc.

[4] I have long urged that the adjective *lingual* be used in the sense of 'pertaining to language', and that the form *linguistic* be limited to the sense of 'pertaining to linguistics' only. Languages and the science *about* languages are on logically different levels! Cf. my note on this in *General linguistics* 2 (1957):42.

[5] Malkiel concluded that the shift from Span. *amplo* (cf. Port. *amplo!*) to *amplio* was probably motivated not only by the frequency of *-ĭdu* as an adjective ending in Latin but also by the fact that the comparative form of Lat. *amplu*, viz. *amplior/amplius*, had given rise to a Latin verb *ampliāre;* and additionally there was the analogic force of such Spanish words as *medio*, *necio*.

[6] It has been said that, while biologists are used to the idea of multiple causation, social scientists tend to oversimplify, in either-or terms, Cf. Guy Richards: ''Sociobiology redefined'', *The Humanist*, Jan.–Feb. 1982, p.31.

REFERENCES

Bueno, Francisco da Silveira. 1967. *Grande dicionário etimológico-prosódico da língua portuguesa.* Vol. VIII. São Paulo: Saraiva.

Malkiel, Yakov. 1967a. ''Linguistics as a genetic science.'' *Language* 43:223–45; esp. 235–43.

———. 1967b. ''Multiple versus simple causation in linguistic change.'' *To honor Roman Jakobson*, The Hague: Mouton. 2:1228–46.

Michaëlis de Vasconcelos, Carolina. *A saudade portuguesa.* 2nd ed. 1922. Porto: Renascença Portuguesa.

Nascentes, Antenor. 1955. *Dicionário etimológico da língua portuguêsa.* Vol. I. Rio de Janeiro.

———. 1966. *Dicionário etimológico resumido.* Rio de Janeiro (?): Instituto Nacional do Livro.

———. 1967. *Dicionário da língua portuguêsa.* Vol. IV. Rio de Janeiro: Academia Brasileira de Letras.

Pap, Leo. 1956. Review of *O falar de Rio de Onor*, by Jorge Dias and J. Herculano de Carvalho (separata from *Biblos* 30, Coimbra 1955). *Language* 32:738–43.

———. 1982. ''The etymology of English *tip* 'gratuity': A case of multiple causation?'' *Forum linguisticum* 6:258–64.

Wescott, Roger W. 1979. ''Lexical polygenesis: words as resultants of multiple linguistic pressures.'' *The fifth LACUS forum 1978*, Columbia, S.C.: Hornbeam Press. Pp. 81–92.

REVIEWS

HUMPHREY TONKIN and KAREN JOHNSON-WEINER. eds. 1986. *Overcoming language barriers: The human machine relationship. Report of the Fourth Annual Conference of the Centre for Research and Documentation on World Language Problems.* New York, December 13–14, 1985.

Reviewed by ERIC WHEELER

With the equivalent of 2,000 books annually to translate into French, English, and a handful of other languages, the Canadian Department of the Secretary of State has good reason to look for computer assistance in a task that has required and still requires intensive human effort. Yet the most successful machine translation project the Canadian delegates to this international conference session on the technology of translation can report is METEO, used since the late 1970's to translate the highly stylized language of English weather forecasts into French. "We like to think of it as being the only . . . fully automatic high quality translation . . . system in existence today, although some 10 to 15% of what it produces requires some human post-editing." (38)

This session of the conference on world language problems was a first-hand account of how those responsible for translation wrestle with the computer technology of today (The other session, reported in a separate volume, is on "The Idea of a Universal Language"; the consequence is that every paper here has an abstract in Esperanto). It has not been easy to come to terms with the technology.

The inability to print the orthography of non-English languages is familiar to practicing linguists everywhere. The extreme case, perhaps, is Chinese with its minimal requirement for 2400 distinct characters—although "only some 558 characters are needed to account for 80% of written texts . . ." (67). The technology is there, with not only large keyboards but also various coding systems that permit character entry from standard keyboards. Yet the problem remains, as much as anything because of the multiplicity of possibilities and the lack of consensus on which system to use. The authors (Wu and Hille of the United Nations translation service) express the hope that "the Chinese people . . . take the first steps in the long journey towards the alphabetization of the Chinese language." (73). Yet, they recognize that the problems are intricately woven into Chinese geography and history, and that the solution to the international translator's problems with Chinese is not simply a matter of more sophisticated computing.

The terminologist (responsible for consistent and definitive translation of expressions) and the translator have fared better from the introduction of word processing equipment and online databases. Like information workers everywhere, the immediate availability of accurate and consistent information, at the work station, has greatly improved the productivity and quality of their work. But it is the cross-linguistic boundaries that frustrate the greater use of the technology—the database that cannot handle the combined features of English, French and Spanish, or the keyboard that requires multiple key strokes to generate characters with diacritics. Clearly, these authors believe it is worth the effort to try to overcome the multitude of such problems that appear. Yet, looking at their efforts to automate from the perspective of 5 years later, it is clear that they had barely begun to implement useful office systems—and one is left to wonder how they could persevere in pioneering systems that seem to generate so many frustrations and produce such limited results, compared to what one could hope for.

One reported experiment that does seem to have met with success was the use of satellites to link remote international conferences to the UN translators based in New York. Using high-grade audio and full-motion video connections, the translators were able to provide immediate translation of the sessions into several languages without having to go to the meeting sites (one in Buenes Aires, the other in Vienna). The translators saw this technology as expanding the demand for their services, and the conference participants were unaware that the translation service was not done locally. Even so, there were adjustments to be made—in the bandwidth of the audio channel to the lighting in the translator's working room. Changes in technology do not come easily.

But, it was human rather than technical challenges that caused the difficulties. One pair of authors (Bowen and Bowen) complain about the misleading expectations created for the technology "when disapproval and ignorance with regard to (human) translation are combined with an unchecked optimism for machine translation . . ." (20)

In retrospect, the problems of keyboards, bandwidth and cheap hardware are all problems that have gone or will soon go away. The problems of people who do not understand that the technology only provides a new method for doing an old job, of people who fear the change or naively believe it will do the impossible, or of people who by their sheer numbers and force of habit cannot change rapidly—these are the problems that cause resistance to an even faster plunge into the ways of the future and that any given project to implement new technology a risky and troublesome venture even in this age of "information technology". For readers who would like to have first-hand experience reports of the impact of technology on language use, I recommend you look into this soft-cover conference report. Although the information on computers and machine translation is dated (how could it be otherwise?), the comment it makes on people and change is timeless.

IBM Canada Ltd., Laboratory
33 Peter St.
Markham, Ont. L3P 2A5
Canada

THOMAS M. STEPHENS. *Dictionary of Latin American Racial and Ethnic Terminology*. Gainesville: University of Florida Press, 1989. 400 + xi pp.

Reviewed by STEPHEN O. MURRAY

This volume contains an enormous amount of information attesting where and when various Spanish and Portuguese lexemes have been applied to a "raza." Like "raza" itself, (see Pitt-Rivers 1973) the terms are often polysemous, highly variable across time and space, and even within a single time and place. To take one of the most heavily used terms as an example, in Peru, "criollo" is glossed as "pure white, born in America" in one source, as "black", the "offspring of persons who are mostly black", "Malaysian", and "native to Lima'' in others (pp. 86–7)—and only the Malaysian report dates from before the 1950s. Similarly, in Brazilian Portuguese, ''crioulo'' varies from ''pure white'' to ''black born in Africa'' or ''black born in America'' (pp. 292–3).

Stephens does not allude to any semantic theory nor to any theory of language change, and despite his meticulous documentation, the ''method'' of compiling these two lexicons—especially the Spanish one—makes the domain appear even more muddled than it is. What term contrasts with what term in what context may still have unexplained variance from speaker to speaker, but native speakers do distinguish persons with some degree of reliability. The natives do not consider the terms as meaningless, and they use (some of) them to classify persons. As Kay (1978:86) noted (in contrasting Tahitian and Brazilian data), ethnic/racial labels ''may be used not only to designate a member of the class of people having a value within a certain range on the continuum corresponding to that word, but also to indicate a contrast between two individuals whose absolute values fall within the same region.'' For non-prototypical instances, distinctions are context-dependent/stimulus-dependent. As for linguists, contrasts are more important for those identifying race/ethnicity, than absolute values of differentiation. Unfortunately, the underlying schema are for all practical purposes unrecoverable from Stephen's compilation. Contrast sets can only be inferred with difficulty (as is also the case in many of the original sources he used). Changes in usage for specific terms are more readily apparent, at least between ''colonial'' and contemporary ones (the vagueness of dating what is meant by ''colonial'' is also in the original sources Stephens consulted).

As Stephens notes, ''each entry and the dictionaries themselves are by nature incomplete'' (p. 9). I am puzzled why ''Zapoteco'' is included, but such other indigenous groupings prominent in the official patrimony of Mexico as ''Azteca'', ''Maya'', and ''Yaqui'' are not. At least the last two name distinct contemporary groupings as well as historically important ones. I am also surprised at the failure to note ''tonto'' in contrast with ''correcto,'' since these are the ''master identities'' in Robert Redfield's (1930) pioneering peasant ethnography. Others who look at the book will (inevitably) have their own cavils.

Despite his recognition that ''racial'' terms are more ''about'' degree of participation in European culture (hence, class) than about physiognomy, I think that Stephen's own implicit schema of ''race'' made him especially attentive to terms for gradations of color. However, since these predominate in the Portuguese half more than in the Spanish half, this may indicate a real difference between Brazil and Spanish-speaking societies (especially those without Caribbean coasts) rather than Stephen's selective perception.

Although I found the book fascinating and erudite, it is hard for me to see how it can be used to build theories of language or of culture, or even descriptions of Brazilian or Spanish-American organization of the domain of "race/ethnicity."

El Instituto Obregón
1360 De Haro Street
San Francisco
California 94107-3239

REFERENCES

Kay, Paul. 1978. "Tahitian words for 'class' and 'race.'" *Publications de la Societé des Oceanistes* 39:81–91.

Pitt-Rivers Julian. 1973. "Race in Latin America." *Archives Européenes de Sociologie* 14:3–31.

Redfield, Robert. 1930. *Tepoztlán*. Chicago: University of Chicago Press.

AUGUST DAUSES. *Theorien des Sprachwandels: Eine kritische Übersicht*. Stuttgart: Steiner, 1990, 93 pp. DM 28.00.

Reviewed by RAIMO ANTTILA

Apart from front technicalities (1–9), the bibliography (91–92), and a term index (93), the book divides into eight chapters: 1. "The physical interpretation: sound law and its consequences" (11–20), 2. "Language change and function" (21–33), 3. "Structure and system" (35–50), 4. "Analogy, frequency and markedness" (51–58), 5. "Simplification and economy" (59–70), 6. "Language change and language consciousness: about idealism" (71–80), 7. "For exemplification: the subject pronoun in Northern Romance" (81–86), and 8. "Summary: philology as descriptive and interpretive science" (87–90). Dauses has other similar succinct little books on grammar, phonology, lexicon, and comparative Romance to his credit (1985–89), all but the last one having *Grundbegriffe* in their titles. As chapter 7 shows here, the material used is again Romance, and the ones whose ears he tries to pin down are likewise colleagues working with Romance (mainly Lüdtke, Mańczak, Martinet and Mayerthaler). The contents listed show in a way traditional concerns, particularly for Romanists, and one might ask whether there is reason for this kind of discussion today. The topic of course fits in his own series of books, but even for the field at large there is indeed ample reason. The main lines are the following:

Most "linguists" would perhaps immediately discard Dauses because he speaks of the philologist, and only about twice adds the term linguist to it (the term, as well as philology itself, is so self-evident that it does not appear in the index). But Dauses is right: in any historical context the philologist is of a higher theoretical order than a "linguist". His steady castigation of philologists who practise armchair analysis (19, 33, 48), or who have been carried away by the system and their own "theory" (17, 39, 47, 57, 87) could be much more caustic, if the reference were theoretical linguists (he does properly point out that a naive generative model is pointless; 56, 67–68). Dauses adds that his criticism is not meant as a total condemnation of the persons criticized, and neither does he exempt himself from the target, and thus what might sound as arrogance is just irony (7). He adds further that his treatment is a kind of primer of philological mistakes (7).

Dauses' constant criticisms are that mechanism and the causality connected with it do not work (25, 30), that there is danger in confusing cause and effect (76, 82, 84), and that description is not explanation (28, 35, 37, 46, 82; too bad, but that is what language is). The search for system-immanent causality is particularly strong in Romance (42), because of the presence of Latin behind it all (89). But language is not closed enough for such an approach (87). Instead, language is primarily a social phenomenon (22, 89), and this goes against the essence of immanent-functional laws (27); the central position in change rests on the flexibility of the speaker (28–29, 49, 85 87). Human agency and other historical factors all have to be considered and woven in, i.e., contexts and situations (21). And the main purpose of language use is to be understood without undue complications (19, 21, 23), and for this the speaker often overdoes it to catch the interlocutor's attention and favor (32; these points come out forcefully also in recent work by Rudi Keller). All this is absolutely right—and all this is not clear to most linguists.

Curiously, word-centered etymology makes it look as though language created holes in patterns which then would be filled out, but this is not the case (49–50). In grammar the danger of taking description as explanation is greater than in phonology, and Dauses lists ten points against predetermination (40–42). Rather than operate with highly abstract metalinguistic categories, the speaker must interpret similarity in context according to his/her pragmatic needs. In other words, analogy will do (43, 45), and happily linguistics is slowly turning to this position, whether helped by research in artificial intelligence or not. Thus we do not need elaborate positional laws or typologically consistent types (42, 43). Formal analogies are social phenomena (53), and since these are general hints rather than laws, they cannot be predicted—rather, analogies are essentially associative (54). This is all very good, and beyond what one gets in "theoretical" linguistics. Linguists are generally totally unaware of the indexical or inductive anchoring of analogy.

Dauses properly dispatches mechanism, but he wants to do the same with teleology (20, 22, 26, 27, 36, 44, 45–46—where he says that analogy works better than teleology, 64, 68), and this is in fact quite prevalent among linguists. There are 6 references to teleology in the index, but there should be 15 more, which shows that the issue is central, more so than causality, which is easy to discard. But here Dauses goes too far. When we talk about human agency and free will, i.e., speakers who will certain events to fulfill their purposes, we are of course within the realm of teleology and final causes. Remember also Peirce's definition of the symbol (and all words are that) as essentially a purpose, not to speak of the practical and abductive syllogisms as necessary in history, and we see that teleology is in fact central. Dauses is not too far off, though, because he does see the difference between goals and consequences (65; cf. Keller's work again) and that between a lawful result and a possibility (85). Whether people pursue their goals consciously or unconsciously is not that crucial, since both factors are clearly there. Dauses deserves credit for daring to mention terms like 'consciousness' and 'idealism' in his title for §7. These have been shunned like poison by linguists, but concepts like these in fact require some kind of purpose, finality, or teleology. They are also making their presence felt in the most recent literature on language origins.

There is much more. The book's pages are dense, and they could easily have been printed out as up to 150, rather than under 100. Dauses discusses the classical problem of push and pull changes (46–49), the distinctive-feature function of phonemes (19–21), the physical non-existence of speech sounds (47), statistics and frequency as solid factors of change (51, 65), the one-to-one relation between form and meaning as a fiction (66, 68, 69), etc. One should not overdo the importance of grammatical signs; they are just useful even when not necessary (29–31). Jespersen's idea of change as progress comes to light in the context of grammatical change as simplification (63). This is one of the spots where ''linguists'' would fare worse, cf. this with the upside-down ''generative'' idea of underlying regularity and simplicity, an idea that matches the Greek notion of pathology: by abolishing the changes, one reached the pristine healthy state of affairs as a starting point (which then, of course, makes rewrite rules akin to ''sickness'').

Even if one does not always agree all the way with Dauses, his book is stimulating and well worth reading. For its size it is indeed rich.

Department of Linguistics
U.C.L.A.
Los Angeles, CA 90024-1543

JOHN VICTOR SINGLER, ed. *Pidgin and creole Tense-Mood-Aspect systems.* (Creole Language Library 6.) Amsterdam: John Benjamins Publishing Co., 1990. xvi, 240 pp.

Reviewed by ROBERT B. LE PAGE

This volume brings together studies of seven specific languages, a reprinting of Labov's 1971 paper 'On the adequacy of natural languages: I. The development of tense', and a very sensible and balanced introduction by Singler. As he remarks, 'creolists have in general rejected the L(anguage) B(ioprogram) H(ypothesis), at least in its strongest form, as the principal means of accounting for the shared properties of creoles. One of the problems with the hypothesis lies in the implausibility of the version of history that it requires . . . If the LBH becomes untenable as an explanatory device for the shared properties of creoles, then it follows that the exclusion from consideration of all but the most radical creoles is also untenable.' (p.ix)

The pidgins and creoles examined range widely. Singler discusses

Kru Pidgin and English; Fayer, the Nigerian Pidgin English of the 18th and 19th centuries; the creoles analysed are Haitian (by Spears), Papiamentu (Andersen), Berbice Dutch (Robertson), Capeverdean Crioulo (Silva), and as an example of an indigenous African creole, Kituba (by Mufwene). In his 1971 paper Labov had taken most of his examples from Hawaiian Creole and Tok Pisin.

I have always been surprised at the amount of attention generated by Bickerton's LBH, since it was based upon an ad hoc constraint of the class of 'true creoles' by conditions extrapolated from the two cases with which Bickerton happened to be most familiar. However, we must be grateful that it has led, as here, to a good deal of careful beavering among much wider examples.

Labov's 1971 paper antedated Bickerton. Labov had come into contact with creoles shortly before the 1969 Mona conference, at which he gave a paper on 'The notion of system in creole languages' (Hymes 1971, 447–72). The title of the paper reprinted here reflects an earlier concern with the supposed 'inadequacy' of some 'languages' and the strategies adopted by some speakers in multilingual communities to remedy that—among them the development of grammatical strategies for tense in emergent creoles. He discussed the association of language use with intelligence and the downgrading of the 'intelligence' of those who used non-standard grammar in tests. He went on to discuss possible ways of studying the relationship between language and thought. The close diachronic/synchronic study of pidgins and emergent creoles has seemed a promising approach to the relationship of forms to functions as these emerge in novel cultures, and so to possible universals of child language development, since creoles were taken to be first-generation nativizations of pidgins.

There are thus two ideologically-related strains in the present search for grammatical universals among creole grammars: the adequacy of children, and the innateness of grammar. There is a third which tends to be recognized only tangentially by most of the contributors to the TMA debate: the carry-over into it from predominantly Indo-European linguistics of notions as to what constitutes the marking of 'tense' or 'mood' or 'aspect', as to the discreteness of these categories, and the discreteness also of grammatical from lexical and from contextually indicated verb categories.

Andersen throws out a hint or two on this last point in his discussion of Papiamentu Tense-Aspect: 'Separate from such morphologically conveyed aspect is what is called 'lexical aspect' or 'inherent semantic aspect' of verbs and predicates . . . The effect of grammatical aspect is often to override the natural semantic aspect of a verb or predicate' (p63). He goes on to discuss Comrie's (1976) binary categories of inherent aspect, and the notion of telicity drawn from Mourelatos (1981), in relation to the semantics of Papiamentu verbs. A useful discussion follows of the interaction between inherent lexical and grammatically-marked tense-aspect in basilectal Papiamentu. He wants to 'provide a model for treating tense-aspect (and any other linguistic sub-system) within the total discourse context instead of the only other existing model, which is sentence-level semantics and syntax alone' (p90). This is a laudable aim.

Mufwene's analysis is of time reference in the eastern dialect of Kituba, one of his own native languages, lexically based on Kikongo. He is concerned with the 'core morphosyntactic strategies of time reference' (p97) leaving largely to one side 'other, predominantly periphrastic, constructions' used to convey e.g. 'terminative or ingressive constructions of recent past . . . only markers that delimit verbs as either inflections or 'auxiliary verbs' are discussed fairly extensively.' I would like to have been told more about the periphrastic possibilities alongside these. However, his analysis yields food for reflection on the adequacy of some of the traditional terminology. One of his conclusions is that 'Kituba's tense system is not articulated in terms of the same basic irrealis/realis or future/non-future distinction as the Atlantic PC systems. If anything, it is akin to the European past/non-past system, since CONCOMITANT DURATIVE may be used for plans.' (p110)

Spears sets out the 'classic case' preverbal marker systems which, Bickerton claimed, Guyanese shared with Haitian. He then goes on to present evidence for a more subtle dissection which depends upon careful semantic analysis of conversational data. He convincingly demonstrates that the actual HC preverbal marker system differs significantly from Bickerton's 'classic case'.

Silva is another native speaker of a creole, Capeverdean Crioulo. He finds that although Bickerton had provided 'the most comprehensive and direct background' for his study, he had to take issue with the classic TMA pattern on a number of points, including the probabilities of historical development. It is not possible to explain discrepancies in terms of the survival of pidgin features. Here we may take a sideways glance at Philip Baker's 'Column' in *JPCL* 52, 1990, where he points out that, by comparing very early with much later data for some creoles, including Mauritian, it is clear that the TMA forms which support Bickerton in fact only emerged gradually over more than one generation. In the case of Tayo, a French-based creole of New Caledonia, it has been the mother tongue of most of the village at least since 1920 but is only now beginning to develop a TMA system. Silva, dealing with iteratives, seems tempted to ascribe the lack of the prescribed system to a distinction between Portuguese-based and French- or English-based creoles.

Robertson, discussing Berbice Dutch, states 'two vital and related characteristics' of Caribbean creoles. The first is that temporal deixis relates to a zero point established by the speaker, possibly having little to do with the time of the discourse; the second, that time values are not obligatorily marked within the verb system but may be signalled in contextual ways. Attempts to isolate the tense-aspect markers of all creole languages ought to be based on an examination of the entire predicate system and be sensitive also to clues from the context of discourse. The base form of the BD verb can express a wide range of tense and aspect, and is not confined to simple past. In addition, BD uses a mixture of preverbal particles and suffixes. Distinguishing the 'genuinely BD system' is complicated by the intrusion of items from Guyanese Creole English as the use of BD dies out.

Fayer relies upon written diaries and letters from Efik merchants in the 18th and 19th centuries. The use in these of Nigerian Pidgin English yields interesting data. The longest early text is a diary of approximately 13,000 words. Tense in this is normally signalled by initial conjunctions and temporal adverbs. *Did* is the most common auxiliary. Fayer would argue that although this is taken from a standard English past form its primary function is for emphasis or contrast, whereas Singler feels that it serves to reinforce temporal sequencing: 'soon after we walk up to see wee town & Did tak one great guns' (I would suggest that it might be fruitful to compare the use of *do* and *did* in conservative West of England and Irish dialects.) The written NPE makes no distinction between stative and non-stative verbs. The system is basically one in which verbs are unmarked. 'The early data do not support the claim that aspect develops before tense . . . Given the infrequency of marking for mood and aspect, it is not possible to determine if there was any ordering of tense, mood, and aspect markers.' (p200)

Singler's own final chapter deals with the impact of decreolization as seen in Kru Pidgin English, since he wants to extend DeCamp's continuum model and Bickerton's TMA prototype to stable and long-established pidgins. He concentrates on the older of the Kru pidgins, the coastal one, with a largely Kru substrate. Basically it conforms to the classic model for mood and aspect but 'departs dramatically..with regard to tense' (p206). Among his five informants, one is at a much more advanced stage of decreolization, as measured by the incidence in his discourse of different non-punctual verb markers. As to tense, the occasional marking that does occur is almost all the result of decreolization; in general KPE is tenseless. It is not possible from today's data to judge whether this was always so.

How much remains today of the evidence for the Language Bioprogram Hypothesis is hard to judge. If we remove the ad hoc constraints on what constitutes a true creole (and, I would argue, if we polish up our knowledge of conservative English and French rural and urban dialects) perhaps not a lot. But we are left with a great deal of valuable discussion about the grammars of a wide variety of contact languages, of which this book is a good example.

University of York
Heslington
York YO1 5DD
ENGLAND

REFERENCES

Comrie, Bernard. 1976. *Aspect.* Cambridge: Cambridge University Press.

Hymes, Dell ed. 1971. *Pidginization and creolization of languages.* Cambridge: Cambridge University Press.

JCPL *Journal of Pidgin and Creole languages.*

Mourelatos, A.P. 1951. 'Events, processes, and states' in P.J. Tedeschi ed. *Tense and aspect.* New York: Academic Press, 191–212.

Cognitive Linguistics Volume 1-1, and *Rivista di Linguistica* Volume 1-1.

Reviewed by JOHN HEWSON

It is reassuring, in these penurious times when all budgets seem to be cut to the bone, that new worthwhile reviews continue to appear. Since such launchings are entirely dependent upon funding and the support of the scholarly community, those who embark on such enterprises are to be commended for their willingness to dedicate themselves to such an arduous task. Setting forth upon the sea when tides and winds are perilous requires navigators of special quality.

It is a pleasure, therefore, to welcome two new international journals, one of which aims to be a forum for cognitive approaches to language and linguistics, and the other to be open to all the diverse streams of contemporary linguistic research, but with a hope of working towards consensus and convergence: the *Rivista di Linguistica* will be geographically based in Italy, but is to be international in outlook and scope. In fact the only item written in Italian in the initial volume is the five page obituary at the end of Giorgio Cardona (1943–1988), one of the original planners of the journal.

Cognitive Linguistics is the journal of the Cognitive Linguistics Association, which was formed in the late 1980's. A basic tenet of the cognitive approach is that the ''formal structures of language are studied not as if they were autonomous, but as reflections of general conceptual organization . . .'' (Editorial statement, p.1). To put it bluntly, grammar (which is too often reduced to mere syntax) is seen to be dependent on semantics; the basic role of the morphosyntax is to mark meanings. It follows that the separation of ''syntax and semantics'' is a fundamental error of method, based on a metaphysical belief that only the directly observable ''tip of the iceberg'' (the ''formal structures'') is truly an object of science. This is the old positivist ideology that

> Genuine will only be
> Things that you can touch or see

a somewhat naive belief that the existence of what is directly observable is somehow beyond question, whereas what is inferrable is ''abstract'' (whatever sense there is to that) or ''exists only in people's heads''. Are we really prepared to believe that gravity, which is what keeps our feet on the ground, exists only in people's heads, is purely imaginary? (If gravity were directly observable, mankind would have had no need of a Newton).

Are we looking at a new intellectual ''revolution'' in this approach? Not really; it is more of a return to a proper understanding of the Saussurian term *valeur,* and of the principles laid out in Chapter Three of the *Cours de linguistique générale* (pp. 114ff), which begins with the subheading ''Duality of all sciences that operate with values''. But there is, on the other hand, a new consensus among psychologists, philosophers, and linguists that language is not to be isolated from other human faculties, and the consequent development of an interdisciplinary Cognitive Science that is very broad in its scope, and that allows the workers in Cognitive Linguistics to take a variety of different avenues, some of which are spelled out in the Editorial Statement, and others illustrated by the contributions to this first volume, which are as follows:

1. Ronald W. Langacker, Subjectification
2. George Lakoff, The invariance hypothesis: is abstract reason based on image-schemas?
3. Wolfgang U. Dressler, The cognitive perspective of ''naturalist'' linguistic models
4. Anna Wierzbicka, The meaning of color terms: semantics, culture, and cognition
5. Gilles Fauconnier, Domains and connections

Although this is a very wide range of topics, all of these articles are eminently readable. There is also a lot of discussion, with questions raised, discussed, and answered, almost as if the writer were discussing with the reader, as in the following where Wierzbicka raises the question: isn't it obvious what *red* and *blue* mean?

> ''No, it is not obvious. It is not obvious at all, and although the psychological, anthropological, and linguistic literature on color terms is very extensive, it usually addresses itself to other questions. The simple, ''naive'' question raised here tends to be largely overlooked—as simple and ''naive'' questions concerning our everyday experience often are.''

The writers not only take pains to make their positions clear; they also, on occasions, take pains to make clear how they differ from others, and even from one another. Lakoff explicitly states the need for this, and shows clearly how he disagrees with Wierzbicka (p.46), having singled her out because of his ''great respect for her distinguished contributions to linguistics over a long career, and because I have learned much from her work despite our disagreements.''

If the editors can maintain this breadth of scope, refreshing directness and lack of jargon, they will have no difficulty finding an audience.

The *Rivista di Linguistica* turns out to be equally broad in its scope, as the comments of the *Sommario* may indicate:

1. Hans Basbøll, Phonological weight and Italian raddoppiamento fonosin-tattico
2. Pierangiolo Berrettoni, An idol of the school: the aspectual theory of the Stoics
3. Pier Marco Bertinetto & Carol A. Fowler, On sensitivity to durational modifications in Italian and English
4. Christina Delogu, The morphological lexicon of a speech recognition system for Italian
5. Alessandra Giorgi & Giuseppe Longobardi, Typology and noun phrases
6. Nunzio La Fauci & Michele Lopocaro, Passifs, avancements de l'objet indirect et formes verbales périphrastiques dans le dialecte d'Altamura (Pouilles)
7. Giorgio Moro, Counterfactual presuppositions of corner conditionals: a re-examination
8. In Memoriam, Alberto M. Mioni, Giorgio Raimundo Cardona (1943–1988)

The first of these is a very interesting discussion of the role of phonological weight at word boundaries, in relation to Italian raddoppiamento and Danish stød, whereas the second is a long historiographic essay on different views over the centuries of Ancient Greek verbal aspect. (This journal also prints in the Greek alphabet, a blessed relief to those who read Greek!) The third is in the psychological side of experimental and acoustic phonetics, the fourth is in the area of Artificial Intelligence, and the fifth in Typology from a Universal Grammar perspective. The sixth is set in a Relational Grammar framework and the seventh in the philosophy of language.

Perhaps the scope here is in fact too wide, since there is very little sign of the "contrapositions" (counterpositions?) or convergences that the editors are looking for, and the last three items in particular are quite narrow in their perspective and unreadable without a knowledge of the specialist jargon that is used by these particular approaches.

The Foreword promises us, however, that every second or third issue will be monothematic and that these issues will focus on highly debated topics. We look forward to some interesting discussions.

Department of Linguistics
Memorial University of Newfoundland
St. John's, Newfoundland
CANADA A1B 3X9

R. HASAN and J.R. MARTIN (eds), *Language Development: Learning Language, Learning Culture. Meaning and Choice in Language: Studies for Michael Halliday* (= *Advances in Discourse Processes* vol. XXVII). Ablex, Norwood, 1989. Pp. ix + 397

Reviewed by NIGEL LOVE

This is the first of three projected volumes intended as both exegesis and celebration of the work of M.A.K. Halliday. A short editorial introduction precedes seven long chapters which explain and extend Hallidayan ideas about language learning and their application to language teaching.

According to the editors Halliday sees linguistics, in the last analysis, as a form of political action. For Halliday a linguistic theory must above all be useful. It should have practical implications and consequences, for instance in the field of linguistic pedagogy. It is not surprising, therefore, that his account of first-language 'acquisition' (a Chomskyan term that Hallidayans reject in favour of 'development') should be continuous with his recommendations as to good practice in the language teacher's classroom. A technically detailed account of the relevance of Halliday's grammatical theories to, specifically, second-language teaching is given in 'Systemic-functional linguistics and its application to the TESOL curriculum' by J. Jones, S. Gollin, H. Drury and D. Economou; but the general idea that linguistics is nothing if not applicable, as it were, is a recurrent theme of the other contributions.

However, an understanding of how the child develops linguistically is not of interest solely to the language teacher. In Halliday's view, to learn anything is fundamentally to undergo or engage in a linguistic experience. This aspect of Hallidayan thought is stressed by D. Butt, in 'The object of language'. Butt's main point is that it is misleading to postulate a simple cleavage between a reified language, on the one hand, which, when mastered, serves as a tool for interpreting the world 'out there', on the other. Rather, there is no world out there, in a sense, except the one mediated for us by language. Learning a language and learning a culture, if not one and the same thing, are at any rate symbiotically interlinked. 'Our external world has to be negotiated, like all our experience, through the semiotic filter of cultural salience'. Not only is there no purely linguistic dimension of experience, but 'there is no activity outside the semiotic organization of the culture, since even the delimitation of events and objects must be deferred to the social order' (p. 109).

C. Cloran's contribution, 'Learning through language: the social construction of gender', offers a particular illustration of the point that learning language and learning a culture are two aspects of a unitary process. Cloran suggests—what is doubtless true—that all cultures build an edifice of social beliefs and practices on the fact that human beings are classifiable in terms of two sexes; and argues that the particular social implications of that classification for a particular culture are both embedded in, and learned through, sex-differentiated patterns of language use.

Given this general perspective, we can see that in Halliday's view an adequate account of language development is required as the basis for an adequate theory of learning itself. On a less abstract level, it has important implications for all formal educational practice. F. Christie's chapter, 'Language development in education', precisifies and elaborates Butt's point. She argues that our culture is strongly inclined to the educationally harmful misconception that proficiency in a language is a mere instrumental skill whose acquisition 'in itself' is a prerequisite for the business of doing things with it, such as learning the subjects on the school curriculum—which, in their turn, are treated as inert 'products' for whose conveyance to the pupil's mind language acts as a neutral vehicle. Such an attitude, argues Christie, ignores two facts: (i) that language is first and foremost a semiotic device, learnt for the purpose of making and expressing meanings, and through the process of attempting to make and express them; and (ii) that a school 'subject' is primarily a particular kind of discourse, constructed out of the characteristic linguistic patterns that embody it. To learn e.g. chemistry is therefore not a matter of first learning e.g. English, and thereby being enabled to acquire a body of information labelled 'chemistry'. It is, rather, to learn to speak and write English in a particular, chemist-like fashion:

> To learn a school subject is to learn how to deal with experience in a particular way. It is to learn to operate as a social scientist, a mathematician, a natural scientist, or as a writer of literature, for example, where each of these involves learning to adopt the appropriate behaviour, including in particular language behaviour (p. 165).

This is not to suggest that, in the early stages of schooling at least, practice or training in the use of language by and for itself can be dispensed with entirely. But such practice should be injected from the outset with significant semiotic content. Language not being a function-neutral skill or facility, language exercises themselves should have an overt, extralinguistic purpose. If the aim of, say, a primary school composition task is to develop the ability to organise a coherent narrative, there is no reason why this should not be made clear to the child, as the explicit object of an otherwise pointless exercise. Two contributions stress the importance of introducing such a metalinguistic dimension into early language training. J. Pearce, G. Thornton and D. Mackay, in their report on a project conducted at University College London designed to bring linguistics into the language classroom, talk of the need for a 'fusion of doing and discussing'; while J. Rothery, in 'Learning about language', recounts the success that was had in explicitly teaching young children to construct written narratives in terms of the tripartite schema 'start', 'complication' and 'resolution'.

The value of these ideas depends heavily on the cogency of Halliday's account of language development *stricto sensu*—i.e. the story he tells about the passage from neonatal languagelessness to becoming a competent participator in a linguistic community. Crushed into the tightest of nutshells, and with apologies in advance for the distortions thereby introduced, this story, as outlined for readers of this volume by C. Painter in 'Learning language: a functional view of language development', may be set out as follows. The starting point is a conviction that how a child develops language is not separable from why. What built-in predispositions cause the child to embark on this process? In sharp contrast with Chomsky's postulation of a genetically given 'language faculty', Halliday starts from nothing more contentious than the suggestion that infants are innately motivated to explore both the persons and the things around them. The neonate's exploration of the human element in the environment takes the form of close attention to people communicating with it; and this attention is manifest in facial, vocal and bodily movements directed towards the communicating partner, who in turn reciprocates the attention by continuously reflecting back to the infant his own communicative moves. The effect is that after a few months the two have achieved a pattern of interaction in which the vocalisations and gestures produced by each participant are synchronised with the other's response. An innate desire for intersubjectivity—a predisposition to attend to other human beings—thus develops through social contact into a communicative game with an adult partner.

This is admittedly a rudimentary form of communication: there are no 'meanings' being 'transmitted' here; we have no more than a synchronised co-operative exchange of attention. But once this has been achieved a new possibility arises for the baby. He is no longer obliged to continue his investigation of the world of persons separately from his parallel investigation of non-human objects: he now has the possibility of using communication with another person to help him to explore the world of objects. An infant may now try addressing another person in order to act on an object in some way—perhaps to have it manipulated, removed, or put within reach—as an alternative to acting on the object directly.

The use of protolanguage allows engagement with a partner in ways that are not open to the entirely prelinguistic infant. But there are severe limitations. The protolinguistic symbols are not the names of things, but express the child's desires or purposes in relation to things. Their meanings are 'functional', not referential. Reference can be achieved at this stage only non-linguistically, by e.g. reaching for the biscuit tin or holding up the toy. This means, among other things, that the protolanguage lacks 'displacement'.

A further handicap at the protolanguage stage is the infant's inability to engage in dialogue in the full adult sense. Protolinguistic 'dialogue' is limited to an initiation of some kind, followed by a response. The same exchange can be endlessly repeated, but conversation cannot be extended by treating a linguistic response from the partner as the basis for a new departure.

Why and how does the child's system develop further? Why should the child wish to do something his protolanguage cannot manage? Because, in developing language in interaction with others, his experience of having his semiotic efforts recognised as such by others (who demonstrate their recognition by responding to his protolinguistic vocalisations) encourages the child to view these others as meaning-makers themselves. This in turn prepares him to attend to the meanings offered by those he addresses, who of course can talk back to him only by using their own system; and so the child's protolanguage changes in a direction entailing the development of *names*.

The difference is that some names are borrowed from adult language rather than invented by the child. The child seems to exploit the fact that the word already meant something to the adult from whom he has borrowed it. Taking over a pre-existing adult word such as *bird* opens up the possibility of initiating a request for a specific object that is not in view and has not been verbally offered.

Again, using the system changes it. In many cases, adult words will at first be used in only one protolinguistic functional meaning, but gradually the child becomes aware of their potential to act as names, as in adult language. We see the child developing the idea of a name when he begins to use the same word for different protolanguage functions.

The child now enters the 'transition phase', in which particular words are no longer wedded to particular functions. When borrowed adult words become names, a child gains tools for learning, for sorting out and categorising experience. And he has come to require this of his symbolic resources not just because of his need to know the world, but because his interactions with others have suggested to him the possibility of language serving this need in this way.

However, the 'functional' nature of the protolanguage as a form of speaker-centered action: to control others, gain goods and services, and intrude upon the world, typically carries over into the transition phase, in as much as although some of the child's words are now names, not all of them are; and there is characteristically a bipartite division of the child's linguistic resources between what Halliday calls 'mathetic' and 'pragmatic' language. Mathetic language is for learning about and reflecting on the world; pragmatic language is for acting on the world.

Many studies have reported that children separate their words into two such classes. Not only do many children develop different versions of their vocabulary items according to which function is being fulfilled, their mathetic and pragmatic linguistic resources may be systematically differentiated, e.g. phonologically: both of the children discussed by Painter distinguished the function of early words by using a falling tone with mathetic utterances, a rising tone with pragmatic ones. Eventually this dichotomy of mathetic and pragmatic verbal systems fades away, as the child comes to realise that both pragmatic and mathetic functions may be combined in any given utterance. Once this has happened, the child's use of language is substantially adult-like.

Painter concludes by drawing out some of the general implications of this account. One is that children learn language in order to achieve something—to understand the environment, to feel close to family members, to obtain desired objects, etc. That in turn implies that language learning is from the outset a matter of making meanings in a particular context of social interaction. Linguistic expressions are not developed for their own sake and then put to use. On the contrary, linguistic expressions (words, grammatical constructions, discourses) are developed in order to actualise meanings.

Despite the particular lacunae and obscurities that different readers will detect in Painter's exposition (exactly why, for instance, should it be impossible to achieve displacement with a child-originated word?), it is clear enough that we have here a welcome relief from, and a formidable challenge to, the aridities of the Chomskyan and other segregationalist accounts of language acquisition, whereby what is acquired is supposed to be a context-neutral system of form-meaning correspondences existing somehow in the abstract irrespective of whether or how it is ever brought into use. As far as this particular collection is concerned, however, the challenge would have been still more formidable were it not for a minor defect and a major omission.

The defect (some may not consider it all that minor) lies in the manner of presentation. The reader of this volume is in for no great literary treat. Every one of these essays is written in that plonking, prolix style so often adopted by the soft-science practitioner seeking the respectability of the hard. Moreover, Halliday is a thinker for whom theoretical and terminological innovation go hand in hand, and who believes that the best way to present complex information is invariably to display it. His followers are faithful to him in these respects, and these pages bristle with neologistically labelled boxes, charts, diagrams, graphs and tables, not to mention long passages of verbatim child-speak. Much of this is necessary, perhaps, but the result is a book less attractive to the uninitiated than it might have been.

What is omitted is any attempt to face up to, explain and, ideally, dispose of the countervailing allure of the Chomskyan orthodoxy in these matters. More than one contributor states that Halliday's approach to child language stands in opposition to Chomsky's, but leaves it at that. Anyone who believes that Halliday's approach is by and large right might be expected to show some curiosity as to why in that case there should exist a large and flourishing fraternity of theorists in this area whose ideas are not only radically different from Halliday's but currently far more influential.

Chomskyans draw a clear distinction between a language and its use. Hallidayans, on the other hand, believe, roughly speaking, that a language is its use. But matters are not so simple: there is no question of a straightforward choice between these positions. Man's linguistic behaviour, like anything else in the world, is liable to fall under man's intellectual scrutiny. Through that process it comes to be decontextualised. Language becomes both a form of action and an object of contemplation. Hence the idea of a language as a context-free system of abstract verbal possibilities, as championed by Chomskyans. But such an idea, or at any rate the germ of it, is an unavoidable concomitant of any attempt we might make to reflect analytically on our linguistic behaviour. And analytic reflection on our linguistic behaviour is no less the business of Hallidayans than it is of Chomskyans. That is, for all their professed belief in language as social interaction, as a process of making contextually determined meanings, Hallidayans, just like anyone else, exploit the possibility of decontextualising language in order to talk about it, for purposes of reporting, illustrating and justifying that very belief. How are we to make anything at all of, for instance, the quoted passages of children's dialogue with which this book is larded, if not on the basis that the language itself is, after all, separable from its use? The fact is that language *is* abstractable from the circumstances of its use; and that abstractability is a precondition of any systematic discourse about it. This is the fact that Chomskyans are relying on, and the undeniability of that fact is one factor in their academico-political success. A truly convincing challenge would involve properly integrating that fact with the no less undeniable fact that language consists primarily in a form of contextually determined behaviour. Let us hope that a future volume in this series will contain the fruits of Hallidayan labour on that integrative task.

Dept. of Linguistics
Univ. of Cape Town
7700 Rondebosch
South Africa

NANCY C. DORIAN, ed., *Investigating Obsolescence. Studies in Language Contraction.* New York: Cambridge University Press, 1989. xi + 445 pp.

Reviewed by LENORA A. TIMM

It is only fitting that a preeminent pioneer in language obsolescence studies—Nancy Dorian—should be the editor of this impressive set of essays devoted exclusively to the topic of contracting languages, for her decades-long documentation of the demise of East Sutherland Gaelic has greatly enhanced our understanding of both linguistic and sociolinguistic aspects of language decline. The geographic scope of the languages reported on in the present volume includes Australia, Southeast Asia, Africa, Europe, and North and MesoAmerica. There are 20 case studies prepared by language-area specialists who, if they have not previously treated the subject of language death per se, bring with them the credibility of long years of contact with their languages and respective speakers.

The book's contributors were asked to respond to four basic sets of questions framed by Dorian in writing their essay: (1) problems in locating terminal speakers and assessing their skills; (2) skewed performance in terminal speech communities; (3) linguistic change and reductive processes as a structured phenomenon; (4) the phenomena of abrupt transition or "tip" and of the persistence against seemingly high odds. Chapters are organized thematically into three sections—the first group focusing on "context", the second on "structure," and the third consisting of invited commentaries from language specialists who consider the implications of research on language obsolescence from their respective disciplinary perspectives. Although the justification for assigning a particular chapter to the first or second section is not always transparent, in general the Section I chapters are more oriented toward providing historical and sociocultural accounts of the decline of languages, while the Section II chapters offer greater detail about the phonological, morphological, and syntactic reflexes of obsolescence. Section III essays serve, each in its own way, as separate reviews of the major contributions to the volume, and will be annotated only briefly in this review.

What do we know about the nature of obsolescing languages: are their structural changes unique or do they share aspects of change with non-obsolescing languages and/or with children's language or pidgin languages? Is change within these languages always in the direction of greater simplicity? How does language variation within dying languages compare with variation found in healthy speech communities? What is the correlation between restrictions (or constrictions) in grammatical structure and reduction in the number of speech genres accessible to speakers? Is it possible to distinguish between change in an obsolescing language that is internally motivated vs. externally motivated—i.e., what is the role of interference from a dominant language in promoting change? Do we always know? These are some of the substantive linguistic questions explored within this volume. Other, more sociolinguistic, issues addressed include the relation of ethnicity to language, how to identify, or define, a "speaker" of a dying language, types of language death, the role of modernization and urbanization in language shift and obsolescence phenomena, and the influence of the symbolic value attached to traditional languages both as a life-preserver and as a factor entering into coding choices. Selected chapters are summarized in the remainder of the review.

Gerrit Dimmendaal surveys different language contact situations in eastern Africa with an eye to finding answers to several questions about the role of economic and social factors in causing the "default assignment" of certain languages, about the relative significance of language as a symbol of ethnic identity, and about the sorts of "traces" associated with language shift in an area. After describing how various hunter-gatherer populations have been displaced by more dominant pastoral groups and of language loss through the interplay of pestilence and (subsequent) economic and political factors. Dimmendaal concludes that factors underlying language shift and the regional or areal ascendancy of one language over others are too complex to be accounted for by appeal to sweeping transformations such as urbanization or modernization. Indeed, it is almost certainly the case that language shift has been taking place in eastern Africa for countless generations. The interplay of tribal, kinship, class, and economic relationships may have differing impacts on language maintenance or shift; and ethnicity or group identity is not irrevocably linked to language, a fact attested to in many instances in eastern Africa. Finally, linguistic traces of languages that have been relinquished may well surface in more recently adopted languages—areas of the lexicon are an obvious candidate for such survival, but relics of the former language's phonology or syntax may also persist.

In using as reference points Scottish and Irish Gaelic in the traditional Gaelic-speaking regions of Scotland and Ireland, Seosamh Watson considers the problem of how to identify a "native speaker" of a language that has for long been in retreat. Watson points out that this is not an issue only for the linguist but also for the speech community itself, in which people are no longer sure who "counts" as a native speaker; the linguist may be the recipient of conflicting advice about whom to seek out for linguistic work on the language. In such situations dialect diversification is often great as well, and another obstacle encountered is that of negative attitudes about the local dialect, often viewed by local speakers as base, corrupt, and not worth the effort of studying it. The prestige attached to the written language, moreover, colors people's perceptions of linguistic knowledge/competence, leading them to identify as speakers those who may have some competency in reading Gaelic, but little or no productive capacity in the language. There is a marked difference among the generations in proficiency in Gaelic, in the predictable direction: younger speakers today may not even be familiar with the traditional Gaelic greetings and blessings. Lastly, the author wrestles with the question of internal vs. external motivation to account for some changes in native Gaelic structures (e.g., loss of the genitive case and the generalization of the nominative; failure to make initial consonant mutations)—specifically, has knowledge of English among Gaelic speakers prompted these and other changes? Only additional research will shed light on this issue.

Lois Kuter describes the historical oppression and stigmatization of Breton, the Celtic language spoken in Brittany in western France for over 1500 years. The language has held out tenaciously in spite of overt tactics by the central government for two hundred years to repress it. Nevertheless, its hold is weakening as the Breton population has, in this century, become virtually bilingual and when the school system has traditionally turned its back on the language, indeed, meting out some unusual punishments for students heard speaking Breton on school premises. The domains of use of Breton have become increasingly restricted among the native speaker population (home, café, certain work situations) and French is everywhere seen and heard as the public language. Kuter's analysis, however, trains attention on the political, socio-economic, and psychological symbolism attaching to Breton and to French; she argues that the symbolic values adhering to each language are in opposition, but that not only negative values attach to Breton, as might be expected given the high level of propagandizing against the language that has gone on. Indeed, in the wake of the cultural-linguistic renaissance of the late 60s–70s in Brittany, Breton has assumed much positive value as a symbol of ethnic (Breton/Celtic) identity and a remainder of Brittany's history. She concludes optimistically that there may be just enough positive symbolism to perpetuate the language for some time to come.

Aleya Rouchdy discusses the situation for the Nubian language (Eastern Sudanic) spoken in Egypt. Its speakers suffered a major disruption following the construction of the high dam on the Nile (in the 1960s), forcing resettlement of some of the Nubian population in villages within a 60-mile crescent; others emigrated to cities. Nostalgia for the "old land" has arisen over time, especially among non-urban Nubians. When, in 1977, the government permitted the formation of cooperatives and established three villages on traditional lands, a sense of Nubian ethnicity was rekindled. However, even a heightened sense of ethnicity is not sufficient to maintain a language in the absence of its use as a vehicle for the primary socialization of children. The latter, in turn, depends greatly on language choice by women, and there are considerable differences among Nubian women in their use of Nubian or Arabic: urban women generally speak Arabic to their children, who grow up understanding Nubian but speaking only Arabic. In rural areas, Nubian monolingualism among women used to prevail, but that has changed as educational opportunities for women have increased and with the advent of radio and television once electricity had been installed in the villages. In addition, more women are now practicing Islam, which further promotes the use of Arabic. In general, Rouchdy's study underscores patterns commonly encountered in minority language situations: the traditional tongue is increasingly spoken by elderly people; it is thought by many not to be useful, or not even to constitute a language since it is unwritten; yet speakers may still have a strong emotional bond with the language. "Tip" to the majority language is well underway, but is more pronounced in urban than in rural areas.

Elizabeth Mertz describes the tip to English from Gaelic in two Cape Breton communities in Nova Scotia. She is interested in the fact that these two communities, though distinctive in socio-economic and religious infrastructures (one is an isolated Protestant fishing village and the other an accessible Catholic town with a diverse economic base), experienced the tip to English at about the same time. Mertz argues that "to understand this phenomenon it is necessary to look beyond regional (or larger) social and economic patterns and to consider linguistic-cultural beliefs and ideologies which filtered and translated wider social change at the local level" (p. 104). The change in use of the traditional language happened fairly rapidly—during the 1930s and 1940s when (fully bilingual) parents began speaking chiefly

English to their children, even though they would continue to speak Gaelic with their friends and with older people. In accounting for this clear case of linguistic tip, Mertz rejects a simple acculturation model, arguing that acculturating factors had been in place for a long time; nor was there any outright repression of Gaelic on the part of the government. The tip seems instead to have resulted from an active decision on the part of Gaelic speakers to ''join the modern world'' and actively to reject an identification of their communities as depressed, economic peripheries of Canada. Yet, though tip has been accomplished, speakers may still deploy their traditional language in creative ways. As a now marked code it can be used to ''create the edges of discourse''—i.e., frame a discourse in English with Gaelic openers and closers. Finally, Gaelic may still serve as a boundary marker between ''us'' and ''them''.

Ruth King focuses on variation in Newfoundland French, taking issue with the claim well represented in sociolinguistic literature that linguistic variation is most likely to take place when there is strong socio-economic stratification. Using evidence from the Arcadian French spoken in Newfoundland, King argues that linguistic variation may exist among fully fluent speakers in the absence of socio-economic stratification. Such variation includes the use/non-use of 3rd plural markers on verbs in relative clauses; and cliticization or not of pronouns and whether or not the clitic attaches to the lower or upper verb in *faire* + infinitive constructions. Age of speaker correlates most strongly with linguistic variability, but this is not, in general, associated with socio-economic categories; neither sex of speaker nor locale appears to influence significantly the variation she has observed. King concludes that the variation she has observed does not carry the same social meaning attaching to variation that has been discussed in the context of urban speech varieties. Her findings are thus consistent with some other work on obsolescent speech communities, notably Dorian's on East Sutherland Gaelic, in which it was shown that variation exists in the absence of social stratification and is, further, not attributable to sex of speaker. On the other hand, some studies of *healthy, non-urban* speech communities have shown strong correlations between linguistic and social variables such as kinship group or sex of speaker. It is likely, then, that linguistic variation in dying languages is of a distinctly less meaningful order than it is in viable languages, whether the latter are situated in urban or non-urban contexts—a finding also documented in the present volume by Mougeon & Beniak for Welland French.

Jane Hill examines "The social functions of relativization in obsolescent and non-obsolescent languages," beginning with the observation that the frequency of relative clauses has been noted in various studies of language death; a lower frequency of relativization has also been noted in oral vs. written language in healthy languages. Aiming at a "functional explanation" of all this, Hill draws on Brown & Gilman's power-solidarity framework to argue that because of the high-information, decontextualizing properties of relativization, higher frequencies of the process would be expected where interlocutors are not intimate and lower frequencies where intimacy is assumed. Coding strategies which use high frequencies of relative clauses will be extended to contexts of "talking up" and strategies with low frequencies will be extended to "talking down." What dying languages, pidgin languages, certain working-class registers, and face-to-face speech share, Hill argues (without, unfortunately, offering any evidence), is an assumption of low social distance and low power differential among interlocutors; low relativization rates, then, are to be expected as a strategy of "solidarity coding." Additional explanations are offered to account for the decrease in relativization frequencies: (1) imperfect acquisition of a dying language and a subsequent simplification; (2) reduction of evaluation of a stigmatized language. Both hypotheses were disconfirmed by data from two Uto-Aztecan languages (Cupeño and Nahuatl).

L. Campbell and M. Muntzel provide a survey of 15 obsolescing languages (14 of them from Mexico and Central America), classifying speakers into four categories along a "proficiency continuum" of strong, imperfect, weak, and rememberers (speakers who never were competent speakers, but who heard enough of the language to remember isolated words and phrases). They propose a 4-way categorization of language death situations and their linguistic characteristics: (1) sudden death due to a catastrophe which virtually eliminates all speakers (e.g. Tasmanian) and, by implication, any obsolescing phase; (2) radical death, typically due to severe economic/political repression possibly with genocidal attacks on a population (e.g., the massacre of 25,000 Salvadoran Indians in 1932 destroyed the Lenca and Cacaopera languages and nearly devastated Pipil); (3) gradual death, the type generally written about in this volume, characterized by encroaching diglossia among a bilingual population, as the dominant language gradually takes over more and more domains of use and as the younger generations gradually lose proficiency in the traditional tongue; (4) "bottom-to-top" death (the so-called "latinate" pattern), in which the traditional language gives way in the home and other everyday contexts, surviving only for ritual/religious purposes.

Proposed also is a classification of the types of change that will take place in dying languages. For phonology they suggest: (i) predictable or expected changes—e.g., the semi-speaker will make fewer phonological distinctions than would a fully competent monolingual speaker of the language; will preserve distinctions found in both languages; will hang on longer to distinctions which have a high functional load. (ii) changes of uncertain predictability—e.g., overgeneralization of unmarked features, or its opposite (overgeneralization of marked features); development of variability (e.g., obligatory rules become optional); development of irregularity by overgeneralizing regular rules; modifying or substituting sounds in the traditional language with sounds of the dominant language. Types of morpho-syntactic and stylistic reduction are also considered. The authors caution that it is not always possible to distinguish change due to obsolescence from change due to language contact that has not (yet) threatened a subordinate language, a theme sounded elsewhere in this volume.

Veteran fieldworker E. Hamp gives us "signs of health and death" in several European minority language or dialect situations (including Albanian, Breton, Slovene, and Scottish Gaelic). He looks first at minority languages that are surviving—with or without interference from surrounding languages. Focussing on phonology and phonetics, Hamp observes that convergence, or loss, in the phonology of such "small" languages or dialects may simply reflect the fact of contact and not necessarily indicate language death (cf. Campbell & Munzel above). Indeed, he wryly observes that "the more contact interference we see—within limits—the more adjustment, adaptation, longevity, and health we may expect!" (p. 201). Among the languages that are indisputably giving way, there are those that are dying "without capitulation" (phonologically, that is)—the salient example is Scottish Gaelic—and those that do "capitulate" (=accommodate) phonologically to the dominant language—e.g., Vannetais Breton and Màndres Albanian (in Greece).

Marion Huffines documents linguistic change in two German-speaking groups in Pennsylvania—nonsectarian farmers, on the one hand, and Amish and Mennonite (sectarian) farmers, on the other. All of the sectarians speak German natively, but all are bilingual. The nonsectarians manifest a spectrum of German-proficiency, and all are competent English speakers. Her linguistic focus in this study is the merger of the nominative and accusative and of the accusative and dative cases in Pennsylvania German (the genitive of European German does not exist in PG). In general, the nominative and accusative have been merged in PG for articles and adjectives, and the dative appears to have weakened after certain prepositions and verbal expressions. From the analysis of interviews with 33 sectarians and 19 nonsectarians, an unanticipated pattern emerged, for the speech of the sectarians showed more convergence to English than that of the nonsectarians, who turned out to be more conservative linguistically. Huffines explains this result by appealing to differences in language use among the two groups: sectarians try to maintain their German and not to use English within the community, in contrast with the nonsectarians for whom English and codeswitching are acceptable. Thus, the sectarians end up looking to English as a source of innovation and lexical creations in their German, which has, accordingly, converged more to English than has the German of the nonsectarians. Apparently codeswitching is working here to maintain the infrastructure of the nonsectarians' German. A further significant finding was that age played a role in the incidence of case merger among the nonsectarians (with the younger speakers, in general, producing more case merger), but it played no important role among the sectarians.

Marianne Mithun compares incipient obsolescence in the Cayuga language (Iroquoian) as it is spoken in Ontario, Canada and in the state of Oklahoma. It is no longer being transmitted as a mother tongue in either location, but among the Ontario Cayuga, 375 adults speak the language on a daily basis and use it for ceremonial events; the Oklahoma speakers use it almost exclusively in ceremonial speeches. Mithun reports that the phonology has changed little in either group, but that the morphology is weakening in its productive aspects. She considers the suggestion that obsolescing languages resemble child language in that those aspects of a language that are the most complex tend to be acquired last and lost first. However, since Cayuga is not being learned by children, she looks at children's Mohawk (also Iroquoian), finding both differences and parallels with adult speakers of an obsolescing language: a major difference is that Mohawk children do not exhibit syntactic complexity for a long time, whereas adult Cayuga speakers do, in spite of less than complete control of productive morphology. Among the parallels are that both children and adults avoid long combinations of morphemes and omit certain morphemes, or make substitutions that would not be done under traditional competence. Finally, there is little recourse to the elaborate word-formation processes that used to be salient in the language.

Edith Bavin prefaces her study of lexical and morphological changes in the Australian language Warlpiri by pointing out that the language is not yet dying out and that there are favorable signs for its maintenance. In this Warlpiri has fared better than many other Australian languages, for "over fifty of the languages spoken when the Europeans settled here already died out, a hundred are on the way, and for only about fifty are the children speaking the language" (p. 268). Nevertheless, Warlpiri has not been immune to change emanating from the language contact situation. This is evident in lexical borrowings from English, and in changes in the pronominal system. Borrowings mainly stand for nontraditional concepts, as might be expected in this type of cultural contact. Some go back to the earliest days of encounter, e.g., *puluki* 'cow' (from *bullock*) and *mikiti* 'gun' (from *musket*), and usually such words are incorporated fully into the morphosyntax of the language. An interesting point is that some loanwords are being used when existing words (and concepts) already were available in Warlpiri (this is especially true of child speakers)—e.g., *boomerang, kangaroo,* and *dingo,* which were borrowed into English from other Australian languages. Warlpiri has a complex set of pronouns that incorporate the inclusive/exclusive distinction, cross-reference clitics, and ergative case markers that attach to long forms of the first and second person independent pronouns. This system is being reduced in complexity and is clearly age-graded, with younger speakers making fewer and fewer distinctions. Bavin interprets this, without much justification, as an internally motivated movement toward greater "semantic transparency" and not as a result of contact per se.

Susan Gal gives a thoughtful analysis of lexical innovation and loss in the Hungarian spoken in a bilingual community of Austria. She is dissatisfied with a common explanation for the differences between healthy and obsolescing languages—i.e., that in the latter case children suffer incomplete acquisition of a language to which they are not sufficiently exposed. Though not without validity, this explanation is inadequate for situations in which structures are being lost from registers or genres of the language that are still being used. Concerning lexical innovation, it has often been said of speakers of obsolescent language that their ability to innovate has been greatly diminished. This is puzzling, according to Gal, in light of data from developmental psycholinguists which show that even young children are able to be very creative lexically, an ability that persists into adulthood. Why, then, should this ability be so affected by the fact of speaking an obsolescing language? Gal's Hungarian data, to the contrary, give strong evidence of continued lexical productivity by adults, but to understand the processes involved it is necessary to examine "the social motivation for the simultaneous occurrence of innovation and loss" (pp. 314–315). The abandonment of one language for another is linked to the symbolic values constructed for the language by the minority speakers in response to their position within a political-economic system; such values mediate not only language choice but also internal change in the lexicon. Her detailed analysis shows that "narrow-speakers" (the younger generation who are less fluent in Hungarian) are in fact lexically quite productive, using word-formation processes that are most salient in the speech of the older generations. However, they innovate lexically just where their elders would not because the latter draw on items well-established for them that the younger speakers have been little exposed to. And they innovate because they still participate in social situations in which Hungarian, not German, is the "symbolically correct" choice.

The book's final section offers "outsiders" perspectives: Lise Menn suggests that Slobin's four cognitive operating principles for the acquisition of (first) language—regularity, frequency, perceptual saliency, and redundancy—may be applicable to speakers acquiring a contracting language. Henry Hoenigswald points to the need for greater precision in evaluating several putative characteristics of dying languages—e.g., style reduction and simplification (loss of categories, allomorphic levelling, etc.)—showing that such processes may well represent general patterns of linguistic change, and are not necessarily diagnostic of language obsolescence.

Kathryn Woolard cautions against what amounts to teleological explanations of linguistic change, putting in a plea for "a focus on actors and speakers" (as opposed to strictly linguistic/structural phenomena) in assessing the reasons for language contraction and death or, alternatively, for language maintenance. Suzanne Romaine considers, with appropriate caveats, some possible shared traits of pidgin/creoles and contracting languages, e.g., the difficulty that often arises of defining the speech community, processes of structural simplification and of relexification, the preference for analytic over synthetic syntax, the rapidity with which change takes place, and the symbolic value of such languages (which may be positive in both types of linguistic situations). Roger Andersen develops the analogy between second language learners (his focus is on immigrants) who have restricted access to the target language for cultural and economic reasons and speakers of dying languages who have restricted access because of the diminishment in functions of the ancestral tongue. He suggests how two basic precepts of second language acquisition may be used to account for many of the "simplification" or "interference" phenomena reported on in language contraction studies (including many in this volume).

Overall, this data-rich collection of essays has succeeded in outlining the boundaries of a dynamic and relatively new arena of sociolinguistic research, has posed numerous, sometimes perplexing, questions about the very definition of "a language" or "a speaker," has documented problems and trouble spots to beware of in carrying out research on contracting languages, has offered an inventory of traits—linguistic and social—that one may expect to encounter in such research situations, and has brought into relief both widespread and particular aspects of languages in retreat. I would expect it to become a classic text in this rapidly developing field of interdisciplinary research.

Department of Linguistics
University of California, Davis
Davis, CA 95616-8685

GERALD KNOWLES. *Patterns of spoken English: An introduction to English phonetics.* London: Longman, 1987. Pp. xi + 262.

Reviewed by JOHN ALGEO

This second volume in the series "Learning about Language" under the general editorship of Geoffrey N. Leech and M. H. Short shares the general aim of the series—to demystify linguistics, to deal with the fundamental facts of the discipline in a way that is comprehensible to an intelligent person. The readership aimed at includes beginning students of linguistics and students of other subjects with a need to know about the basics of language.

In accord with that aim, the book covers not merely the traditional subjects dealt with in books on English pronunciation, but much else as well. Traditional matters are not neglected, but they are often combined with subjects of interest to nonlinguists. The first chapter covers general matters of language and variety, speech and writing, and an overview of sounds, but also the phonological analysis of written texts, including homophony and homonymy. Chapter 2 deals with the articulatory mechanisms, but also with the esthetics of sound. Chapter 3 is on the sound units of English and onomatopoeia; chapter 4, on the syllable and rime, alliteration, and allied matters of sound patterning.

Chapters 5–7 deal at far greater length than is customary in general phonetics books with accent—rhythm, word stress, phrasal stress. Chapters 8–9 cover intonation, including its correlation with punctuation; verse units; and phatic expressions of greeting, thanking, apologizing, and so on. Chapter 10 deals with the expression of "attitude" in language through tone of voice—various paralinguistic phenomena—and the phonological correlates of text structure, as well as with the ways authors have of hinting at such features in a written report of conversation.

The refreshingly broad scope of this book should make it highly useful for students of language and notably of literature in the universities of the United Kingdom. In the United States unfortunately, general students of English—that is, most English majors—have very little exposure to the study of the English language. Almost none of them will take a course that uses a phonetics book, even such a broad one with literary implications as this book has.

Moreover, regrettably this book is not suitable for use with even those few American students who take a phonetics course. It is a British book, no doubt excellent in a British context, but its content would provide unnecessary roadblocks for the American student.

The preface announces that the book is not going to be limited to RP but will treat other varieties of English pronunciation. And it does so, at least to some extent. However, the linguistic facts of the United Kingdom being what they are, the only pronunciation treated thoroughly in the book is RP. Other varieties are referred to passim. Most of those references will be useless to American students, for understanding them fully requires that the reader be familiar with those regional accents and often understand their difference from RP as well.

American English is treated (according to the index) in fourteen places in the book. Three of these treatments are general references. One comments on the American pronunciation of *tomato* and *vase* (and gets the final consonant of the latter wrong, the only American pronunciation cited being /veiz/, whereas the usual American form is

/veis/). One is on the phonetic quality of American /æ/; others on that of /ei/ and /ou/, /ai/ and /au/ and the vowel of *caught*. There is mention of the vowels of words like *class* and *of*. There are references to the American pronunciation of initial *h*, of *r* after vowels and of vowels before it; the dropping of the /y/ glide in *tune, duke, news,* and the like; the dark quality of *l; marry/merry/Mary* homophony; stress in words like *matrimony* and *laboratory;* and intonation in tone groups with two stressed syllables. Obviously, several important characteristics have been omitted. While there is some reference to regional variation within Britain, there is practically none to regional variation within America. No introductory book can cover all features of all major dialects. But this book is clearly directed to British students and does not speak to Americans.

A more serious failing—although a very wide-spread one in Europe—is that the book perpetuates the myth of "General American." The concept of "General American" has a reality and a use—but both are restricted and essentially negative. In teaching their national variety of English to foreigners, knowledgeable Americans will avoid any of the features of pronunciation that are strongly suggestive to Americans of regional identity; they will eliminate those features and teach instead ones that are as general as possible. Similarly, if anyone wants to "improve" their speech, to escape regional identification, they will eliminate locally distinctive features from their pronunciation.

That is, for such purposes we use *r*-full pronunciation, not the *r*-less pronunciation of eastern New England, New York City, and the deep South. We use lax, fairly short vowels in *caught* and *bad,* not the long tense vowels of the Northeast, nor the diphthongized vowels of the South. We use the same medial voiced flap in *liter* and *leader,* not a contrast of voiceless-aspirated versus voiced stops. And so on. Such a pronunciation, from which all the most readily recognizable geographical and social markers have been eliminated, may be called "General American."

But this "General American" is not an institutionalized, clearly defined or recognized accent—either geographical or social. It is not a necessity or even a desideratum for political or social advancement in the United States. It is a negative abstraction—what is left after the most easily identifiable features of regional pronunciation are eliminated. It is not at all parallel to RP, which is a highly institutionalized, well-defined, and still very powerful talisman of social standing.

The second edition of the unabridged *Random House* defines *General American* as follows:

> any form of American English speech considered to show few regional peculiarities, usually including all dialects except for eastern New England, New York City, Southern, and South Midland (no longer in technical use).

That is a good definition. The term does not denote "a form" of American, but "any form" that cannot be easily located. And it is not used technically because it does not correspond to any clear referent.

One wishes that Europeans consulted American dictionaries about American matters. Continental Europeans often talk about "General American" as though it were a regional dialect. Britons tend to talk about it as though it were a social dialect parallel to RP, and so this book does:

> It is much easier to generalize about American pronunciation, and the standard model is actually called GENERAL AMERICAN or *GenAm*. We shall be referring to RP and GenAm collectively as the '*standardized varieties*' of English.

There are two things wrong with statements like that. First, there is no "standard model," no "standardized variety" of American pronunciation. There are regionally and socially and educationally marked varieties—but no coherent accent comparable to RP. When Margaret Thatcher wanted to get ahead in the Tory Party, she had to suppress her native Lincolnshire, shopkeeper's daughter accent and cultivate a variety of RP. John Kennedy, Lyndon Johnson, and Jimmy Carter all continued to sound like Boston, Texas, and Georgia respectively. Second, American dialectologists do not use the term "General American" at all, so this imaginary American standard model is actually so called mainly by foreigners. Some American phoneticians and speech-correctionists have used the term, but their use has not been adopted by language scholars generally.

A book mainly about British English written for British students cannot be expected to deal with American pronunciation in detail. However, no book should promulgate error. Most of what is written about "General American" is just that. It is unfortunate that this otherwise generally competent, innovative treatment should perpetuate an ill-informed and misleading idea.

Dept. of English
The University of Georgia
Park Hall 254
Athens, GA 30602

ROBERT A. HALL, JR. *A Life for Language: A Biographical Memoir of Leonard Bloomfield.* Amsterdam & Philadelphia: John Benjamins, 1990. 129 + x pp. $29.00 hb.

Reviewed by STEPHEN O. MURRAY

Hall has crafted a very readable and engaging portrait of a quite reticent man who was fiercely committed to building and applying a science of linguistics to American education and life. This slender volume makes the personality of Leonard Bloomfield (1887–1949) clear. It is not a thorough intellectual biography such as Regna Darnell's (1990) Sapir biography, but manages to give the reader a definite sense of what Bloomfield was like, ''blind spots'' and all. Hall is intimately familiar with Bloomfield's writings and provides recollections of personal contact with Bloomfield. He does not claim to have been an intimate of Bloomfield's. Indeed, from what Hall and others report, it appears that Leonard Bloomfield's only intimate was his wife, née Alice Sayers. Hall portrays her as fiercely protective of her husband and very concerned about any perceived slight of her husband's pre-eminence (also see Esper quoted in Hockett 1970:538). Hall is far too respectful and diplomatic to suggest that she may have scared away or driven off peers and contributed to an isolation toward which Leonard Bloomfield's obviously reclusive temperament tended.

Leonard Bloomfield was not a vigorous self-promoter. However, he could be quite aggressive in insiting on the importance of linguistic science, especially within education, the domain in which he made his living. Hall's most original contribution in this biography is showing that (contrary to the claims of Hockett 1970:543 and Hall 1975:55) Bloomfield was not unappreciated by academic administrators. In particular, using University of Chicago records restricted until 1989, Hall proves beyond any doubt that the University of Chicago (in the person of Humanities Division Dean Richard Patrick McKeon [1900–1985]) considered Bloomfield an important enough asset to attempt to retain him in 1939. At least this reader is curious about the other end of the competition for Bloomfield's services. That is, who pushed the Yale offer of a Sterling professorship to him?

It bears stressing that despite his ascetic refusal to celebrate his own accomplishments, Bloomfield's academic career and professional reputation steadily rose, culminating in the Sterling chair at Yale and being revered by the next generation of linguists—to the point of having his name used as the appellation for the dominant American linguistics at the time of his death (''Bloomfieldianism,'' now more commonly termed ''neo-Bloomfieldianism''). Unfortunately, Mrs. Bloomfield was unable to enjoy—or even notice—this worldly success. The elevation to Yale turned out to be a poisoned reward for Bloomfield, as earlier it had been for Sapir (Murray 1981:66, 1986: 272; Darnell 1990:400-2, 416). Like Sapir, Bloomfield entertained very naive notions about the normal realities of American academic labor (see pp. 30, 41, 64). Neither of them recognized that they were appreciated and well off at Chicago. Neither of them foresaw the problems moving to New Haven was going to create in their lives.[1]

The summer before the call to Yale, Bloomfield had his ''one major triumph as a teacher'' (Hockett 1970:546) at the 1938 Linguistic Institute (LI). Hall (p. 52) notes, ''The 1938 general linguistics course on general linguistics was particularly successful, being attended by virtually all the faculty and students, and arousing great interest and enthusiasm,'' but does not explain what prompted the enthusiasm or describe what Bloomfield taught there. From his own experience as a student at Chicago, Hall reports that Bloomfield ''was a quiet but very firm and demanding teacher, insisting on exactitude in both translation and linguistic analysis'' (p. 42). Hall quotes briefly from published recollections by William Moulton (1970) and Kenneth Pike (1989). He does not seem to have attempted to elicit memoirs of others, and does not cite published recollections by Harris (1972) and Voegelin (1974) which seem to me to add to the thin documentation on how and what Bloomfield taught linguists.

What qualities made Bloomfield an inspiration to the linguists in the wartime Intensive Language Program (ILP) of preparing analyses and teaching materials in potentially militarily strategic languages are not described (or even listed). Like the 1938 LI, what constituted his example to other ILP linguists is left a mystery to be taken on faith. Hall (pp. 75–76) elucidates Bloomfield's role in producing the army's *Spoken Russian* and in inventing a biography for its listed author, I. M. Lesnin, ''an approximation of a Russian pronunciation of 'I am listening,''' but forebears mentioning Bloomfield's dissatisfaction with other work by George L. Trager. Hall provides a very clipped account (p. 74) of how there came to be two manuals of Dutch. Joos (1986: 124) provides a fuller account, that, I think, provides some insight into this very unworldly scholar's somewhat surprising worldly successes. The manuscript of *Spoken Dutch* Bloomfield wrote

> had been extensively (and unsystematically and irregularly) altered by War Department functionaries under a superior officer who had undertaken to en-

> force the right to 'review and adopt' as stipulated in publication contracts . . . This particular Author adamantly refused to permit publication of anything like the page-proofs he had received.

In the context of the winter of 1943/1944, Dutch was a vitally important language. The American military was planning an invasion of northern Europe. Moreover, Dutch was important as a lingua franca in what had been the Dutch East Indies prior to Japanese conquest. And there were military plans involving Dutch Guiana (Surinam) as well. The American military had multiple, pressing needs to train some people to speak Dutch.

> All these enterprises were at the mercy of one cloistered academic in New Haven, Connecticut. What to do? Admirals joked about blowing him out of the water, colonels about dropping paratroopers to take him alive; but then it turned out that he had never been personally under contract for Dutch work and still owned the rights. How about a rush job by somebody else? No, all competent personnel were busy . . . The ACLS [American Council of Learned Societies, which co-ordinated the civilian program] had money to make promises with, yet money seemed to mean nothing to Leonard Bloomfield personally; what he cared about, apparently, was something said to be named 'academic honesty' or the like, and on that basis he drove a hard bargain, to judge from his demands that were met. First off, he required the 'classroom book' [*Colloquial Dutch*] to be scheduled for immediate printing and binding (instead of the originally intended post-war scheduling) and distributed to the same military units that got *Spoken Dutch* (Joos 1986:124).

Not much is recoverable about Bloomfield's early years, but I think that Hall could have told readers more about how Bloomfield worked with informants and with younger linguists both during the wartime endeavors and at Linguistic Institutes. Hall does not even draw on Ives Goddard's (1987) painstakingly detailed reconstruction of Bloomfield's methods of recording, analyzing and standardizing Menomini,[2] although this exceptionally illuminating contrast of Bloomfield's actual methods with the methodology prescribed in *Language* (Bloomfield 1933) was published in a book Hall edited (Hall 1987) and adumbrated by Voegelin's (1974:36–37) recollection of the less-than-totally-inductive method Bloomfield actually used. The slenderness of the volume, thus, is not simply a function of lack of information on which a biographer/historian can draw.

On its own modest terms as "a biographical memoir," it is perhaps inappropriate to complain that Hall did not accomplish more than he set out to do. He has written a concise, discerning and quite moving brief account of Bloomfield's life. Although the story he tells is a tragedy, Hall also shows the saint at play. Indeed, Hall seems to give more space to what Cowan (1987) called ''the whimsical Bloomfield'' than to Bloomfield's linguistic contributions. I wish there was more of the latter, but not that there was less of the former. The parts of the book that contain examples of Bloomfield's elaborate word games, ironic jests and practical jokes not only show that the archetype of American methodological asceticism was human, but show that his aesthetic senses were not undeveloped, as some Sapir admirers have supposed.[3]

Anyone who has actually read Bloomfield (a set not including many linguists trained outside the United States or inside the United States after 1960) knows that he wrote very clearly, wasting no words. He wrote in a clipped, modernist style quite unlike the sometimes hyperbolic, sometimes old-fashioned and meretricious, but often vivid lyricism of Sapir. Apparently, Bloomfield thought that his 1933 book *Language* was written simply enough ''to be understood by the ordinary high school student'' (p. 44). He overestimated the concentration of the ordinary high school student (as he overestimated the ability of professional linguists to understand so concentrated an exemplar of analysis as his 1942 Ilocano paper) and later had to acknowledge that he had not succeeded in explaining linguistics to the general public, however influential his book was on those professionals and apprentices trying to describe languages in the 1940s and 50s.

Those who have been labeled ''neo-Bloomfieldians'' paid too much attention to Bloomfield's synchronic methodology codified in *Language* (1933) and not enough to his method, i.e., what he actually did in describing natural languages. Bloomfield's actual methods involved unobservables (such as the community norm and the standard dialect), and he certainly did not eschew consideration of meaning or even use of comparative data (see Goddard 1987; Voegelin 1974; Fries 1954, 1961; Pike 1989). When they possessed diachronic/comparative data, neither of the fathers of American structuralism performed strictly synchronic analysis (for Sapir and his students, see Murray 1983a; Silverstein 1986; Goddard 1986; Darnell 1990).

Hall (p. 90 and elsewhere, e.g. 1975:58) denies that there was a ''Yale school'' or a ''unified 'Bloomfield school,' '' but recognizes that some linguists of his generation—in particular another methodology mistaken by those who followed it as a method, Harris (1951)—attempted after World War II to banish ''meaning'' from linguistics. Hall was never identified with those to whom he applies the label ''extremist.'' Bloomfield himself was not that kind of ''Bloomfieldian,'' and expressed skepticism about the ultra-rigorism some members of the succeeding generation were attempting (see Bloomfield 1945; Fries 1954, 1961; Koerner 1970; Murray 1983:126–27; Pike 1989). The revolution within American linguistics that bears the name of Noam Chomsky (b. 1928) rejected Bloomfieldian methodology (Harris 1951), and much of Bloomfield's methodology (Bloomfield 1933)[4] often without knowing what Bloomfield did and without being able to do what he did (e.g., use comparative material).

The hardcover Holt edition of Bloomfield's (1933) *Language* remained in print until 1976, and sold more than twice as many copies after the dawn of 1957 than before then.[5] In contrast, until 1970 (with the publication of Hockett 1970) Bloomfield's other writings were not readily available. I would agree with Hall (p. 92) that Bloomfield's reputation was at a low ebb during the 1960s, but I think that serious reconsideration began long before the centenary of Bloomfield's birth in 1987. There were two impassioned books by Erwin A. Esper (1895–1972) stridently defending a behaviorist linguistics (Esper 1968, 1975; cf. Harris 1972:253; Voegelin 1974:36). The first was given an extensive and sympathetic review in *Language* (Olmstead 1970). Most significantly, Hockett's (1970) collection with incisive commentaries on Bloomfield's writing, as Voegelin (1974:35) only somewhat hyperbolically claimed, "may safely be said to provide a turning point in the history of science approach to what has been said and is being said by linguists and about linguists in their work." At the very least, that impressive volume was a turning point in rescuing what Bloomfield wrote from the legends of later structuralists, both those purporting to build on his heritage and those claiming to reject it altogether. Following upon important articles published around the same time as Hockett's anthology by Koerner (1970), Stark (1972) and Silverstein (1973), there were sustained considerations of Bloomfield as a major source of contemporary American linguistics in the lengthy histories of American linguistics by Hymes and Fought (1975:1002–1092) and Murray (1983b:89–151). Both of these works were written years before the publication dates. And both realigned Bloomfield with Sapir in contrast to "neo-Bloomfieldians."[6] Significantly, even Mary Haas (1976:64), one of Sapir's linguistic students least influenced by Bloomfield and who has also exhibited active hostility to neo-Bloomfieldians/-ism, credited Bloomfield rather than Sapir with pioneering fieldwork by immersion in the language to be analyzed. Hall cannot be unaware of the existence of this material published in what he regards as Dark Ages in discourse about Bloomfield between Hall (1950) and Hall (1987), though only Hockett (1970), Koerner (1970) and Silverstein (1973) are in the bibliography, and none of these works is mentioned in regards to the "low ebb" (p. 92).

It seems to me the last chapter (especially the last paragraph) of Hall's book is of a considerably lower quality than the rest of the book. Some of the scholarly energy devoted to preparing detailed name and topic indices should, in my opinion, have been expended on examining the secondary literature if claims about Bloomfield's posthumous reputation are going to be made. The embattled tone (recalling Esper's fulminations) of the last few pages detracts from a book which up through Bloomfield's death is exceptionally fair-minded for a historical memoir.

Despite this unfortunate ending, I think that this volume is essential to anyone interested in the history of linguistics in the United States during the 20th century, not least for the updated list of Bloomfield's publications included. I wish that Hall had striven for a more detailed intellectual biography, but I also want to stress that the brief biographical memoir that he has chosen to write is a moving tale of slowly-building professional triumph, personal inhibition and personal tragedy. Perhaps it is just because Hill brings Bloomfield back to life as a human being (not just the promulgator of a doctrine or the producer of this or that analysis) that the reader (at least this reader) wishes there was more.

El Instituto Obregón
1360 De Haro
San Francisco, CA 94107-3239

ENDNOTES

[1] The anti-semitism of Old Yankee New Haven and intra-departmental rivalry were of less relevance to Bloomfield than they had been to Sapir. Actually, the deep depression Mrs. Bloomfield entered when she was uprooted from Chicago parallels that of the first Mrs. Sapir when she was uprooted from New York three decades earlier (see Murray 1981:66; Darnell 1984:183, 1990:133–36). Nonetheless, leaving Chicago was tragic for both Bloomfield and Sapir.

[2] Hall quotes expressions of Bloomfield's despairing anger at the imposition of American peasant life on the native people of Wisconsin from a 1920 letter (p. 20) and from the preface to the 1928 *Menomini Texts* (quoted, pp. 64–65). I feel that Bloomfield's Americanist work is slighted by Hall. That Bloomfield served on the ACLS Committee on Research in Native American Languages (see Leeds-Hurwitz 1985; Darnell 1990:277–87) is not mentioned, and there is nothing aimed at sorting out the contributions of Hall's Cornell colleague, Charles F. Hockett (b. 1916), from Bloomfield's in the posthumous Algonquian works that Hockett prepared for publication.

[3] Sapir was renowned for his brilliance, but was filled with self-importance, quite unlike Bloomfield. Sapir was quite lacking in a sense of humor on-the-job or off.

[4] But not formalism, nor the idealization of "la langue" (with the concomitant ignoring of "la parole"), nor the autonomy of linguistic analysis from any psychology (even cognitive; see Lamb 1967).

[5] 12,800 copies were sold between 1933 and 1956, 26,800 between 1957 and 1976. Sales information was supplied to me by John Dill of Holt, Rinehart and Winston in a letter dated 26 July 1977. The University of Chicago Press issued an edition of *Language* with a foreword by C. F. Hockett in 1985.

[6] Following Fries (1963), the most important consideration of Bloomfield's work between obituaries and Esper (1968).

REFERENCES

Bloomfield, Leonard. 1928. *Menomini texts*. New York: American Ethnological Society.

———. 1933. *Language*. New York: Henry Holt.

———. 1942. "Outline of Ilocano syntax." *Lg*. 18.193–200.

———. 1945. "On Describing Inflection." *Monatshefte für deutschen Unterricht* 27.8–13.

Cowan, J. Milton. 1987. "The whimsical Bloomfield". 23–38 in Hall (1987).

Darnell, Regna D. 1984. "The Sapir years at the Canadian National Museum in Ottawa". *Edward Sapir: Appraisals of his life and work,* ed. by Konrad Koerner. Amsterdam & Philadelphia: John Benjamins. Pp. 159–168.

———. 1990. *Edward Sapir, linguist, anthropologist, humanist*. Berkeley: Univ. of California Press.

Esper, Erwin A. 1968. *Mentalism and objectivism in linguistics: The sources of Leonard Bloomfield's psychology of language*. New York: Elsevier.

———. 1975. *Analogy and associationism in linguistics and psychology*. Athens: University of Georgia Press.

Fries, Charles C. 1954. "Meaning and linguistic analysis". *Lg*. 30.57–68.

———. 1961. "The Bloomfield school". *Trends in American and European linguistics, 1930–1960,* edited by Christine Mohrmann, Alf Sommerfelt & Joshua Whatmough, 196–224. Utrecht: Spectrum.

Goddard, Ives. 1986. "Sapir's comparative method". *New perspectives in language, culture, and personality,* ed. by William Cowan, Michael Foster & Konrad Koerner, 191–214. Amsterdam & Philadelphia: John Benjamins.

———. 1987. "Leonard Bloomfield's descriptive and comparative studies of Algonquian". 179–217 in Hall (1987).

Haas, Mary R. 1976. "Boas, Sapir and Bloomfield". *American Indian languages and American linguistics,* ed. by Wallace Chafe, 59–69. Lisse: Peter de Ridder.

Hall, Robert A. Jr. 1950. "Obituary of Leonard Bloomfield". *Lingua* 2.117–23. Reprinted in Hockett (1970:547–553).

———. 1975. *Stormy petrel in linguistics*. Ithaca: Spoken Language Services.

———. 1987. *Leonard Bloomfield: Essays on his life and work*. Amsterdam & Philadelphia: John Benjamins.

Harris, Zellig S. 1951. *Methods in structural linguistics*. Chicago: Univ. Chicago Press.

———. 1973. Review of Hockett (1970). *IJAL* 39.252–55.

Hockett, Charles F. 1970. *A Leonard Bloomfield anthology*. Bloomington: Indiana University

Press. (an abridged paperback version was published in 1987 by the University of Chicago Press).

Hymes, Dell H. & John Fought. 1975. "American structuralism". *Current trends in linguistics* 13.903–1156.

Joos, Martin. 1986. *Notes on the development of the Linguistic Society of America, 1924–1950.* Ithaca, NY: Linguistica.

Koerner, E.F.K. 1970. "Bloomfieldian linguistics and the problem of meaning". *Jahrbuch für Amerikastudien* 15.162-83. Reprinted in *Toward a history of linguistics* (Amsterdam & Philadelphia: John Benjamins, 1978. pp. 155–76).

Lamb, Sidney M. 1967. Review of Noam Chomsky, *Current issues in linguistic theory* (Cambridge: MIT Press, 1964) and *Aspects of a theory of syntax* (The Hague: Mouton, 1965). *American Anthropologist* 69.411–15.

Leeds-Hurwitz, Wendy. 1985. "The Committee on Research in Native American Languages". *Proceedings of the American Philosophical Society* 129:129–60.

Li, Fang Kuei. 1988. *Linguistics, East and West.* Berkeley: Regional Oral History Office.

Moulton, William G. 1970. "Bloomfield as Germanist". 512–23 in Hockett (1970).

Murray, Stephen O. 1981. "The 'Canadian winter' of Edward Sapir". *HL* 8.63–68.

———. 1983a. "The creation of linguistic structure". *American Anthropologist* 85.356–62.

———. 1983b. *Group formation in social science.* Edmonton: Linguistic Research, Inc.

———. 1986. "Edward Sapir in the 'Chicago School' of sociology". *New Perspectives in language, culture, and personality,* ed. by William Cowan, Michael Foster & Konrad Koerner. Amsterdam & Philadelphia: John Benjamins. Pp. 241–92.

Olmsted, David. Review of Esper (1968). *Lg.* 46. 131–140.

Pike, Kenneth L. 1989. "Recollections of Bloomfield". *HL* 16.217–25.

Sayers, Frances Clarke. 1987. "The small mythologies of Leonard Bloomfield". 15–21 in Hall (1987).

Silverstein, Michael. 1973. "'Distinctive features' in Bloomfield's phonology". Unpublished Manuscript.

———. 1986. "The diachrony of Sapir's synchronic linguistic description". *New perspectives in language, culture, and personality,* ed. by William Cowan, Michael Foster & Konrad Koerner. Amsterdam & Philadelphia: John Benjamins. Pp. 67–109.

Stark, B. R. 1972. "The Bloomfieldian model". *Lingua* 30.385–421.

Voegelin, C.F. 1974. "Review of Esper (1968), Harris (1973) and Hockett (1970)". *Language sciences* 30.35–38.

DAYLEY, JON P. 1989. Tümpisa (Panamint) Shoshone Dictionary. Linguistics Volume 116. Berkeley: University of California Press. Pp. xxxviii, 516. $56.00.

Reviewed by CHARLES PECK

Dayley's dictionary of Tümpisa Shoshone is small but well constructed. It is small because it includes only 3500 entries. An exhaustive dictionary would probably include around six or eight thousand entries.

Dayley recognizes the limitations of his dictionary. He knows that he would have to live with the people for years to hear all the place names, names of all the plants, animals, insects, medicines, and so on. On the other hand, he has included some rare terminology, such as *pittatahan* 'bare one's behind (of a woman)', which he collected because it was in one of the folk tales he collected. (That is the way it is with the rarer terminology. Over a year or a decade, the frequency of such a term is vanishingly small. But when the occasion to use it comes up, it can be temporarily frequent.)

The dictionary is well constructed because it begins with a good overview of the phonology and orthography, has illustrative sentences on many entries, includes affixes in the dictionary and has an English-Tümpisa Shoshone glossary at the end.

The phonology is complex and the actual pronunciation is not obvious from the orthographic spelling. So the explanation of the phonology and orthography is very welcome.

The entries are fairly complete, with any irregular forms included, with one or more illustrative sentences and frequent cross-references. Usually the gloss is just one English word, which is all right, but probably not true. English, with its hundreds of thousands of words, usually has several words that would be translated by just one Shoshone word.

The inclusion of affixes in the dictionary is welcome—so many dictionary makers leave the affixes in a grammar statement and do not put them in the body of the dictionary. When I tried to parse an illustrative sentence, I could find all the roots and affixes in the dictionary—so parsing is possible with this dictionary.

The English-Tümpisa Shoshone glossary will be useful to people doing comparative work, to people wanting to learn to speak to write the language.

Unfortunately, the language is nearly extinct. Only a few very old people still speak the language. The younger people use English. So this dictionary will be chiefly of academic interest. Dayley has captured the language before it is completely lost, for which we should all be grateful.

JAARS Center
P.O. Box 248
Waxhaw, N.C. 28173

DAYLEY, JON P. 1989. *Tümpisa (Panamint) Shoshone Grammar. Linguistics Volume 115*. Berkeley: University of California Press. Pp. xiv, 522.

Reviewed by CHARLES PECK

The Tümpisa (Panamint) Shoshone language is a member of the Numic sub-grouping of the Uto-Aztecan Family of languages, which were/are spoken by people living in the Great Intermountain Desert from Idaho in the North to far down in Mexico, in the south. The Tümpisa Shoshone live in south-western Nevada and southeastern California. The Name 'Panamint' has been used in the literature but makes no sense to the people, so Dayley does not use it. In Chapter one, 'Introduction,' Dayley discusses the previous literature and describes his own research.

In chapter two, 'Language Typology and Sentence Structure,' Dayley gives an overview of the structure of the language. He describes the agglutinative structure of Shoshone verbs. Verbs often have from two to six affixes, plus an incorporated auxiliary verb. Verbs and verb roots indicate plurality of the object by changes in the verb root or by reduplication (in some verbs).

Word order in clauses is very flexible; since objects carry an object marker, there is no confusion between subjects and objects. More often than not, the object precedes the verb. The subject precedes the verb when it is encoding new information. The subject is absent or is post verbal when it encodes old information, loosely defined. It is also common for a deictic pronoun, such as 'that' or 'those' to occur early in the clause and a fuller expansion of the noun phrase later in the clause, especially when the noun phrase is repeating old information.

The present tense 'verbless' [my term] clauses consist of a noun phrase and a predicate noun phrase, adjective phrase, or an adverb phrase. The language does have a copulative verb *naa* that can be inflected for other tenses and aspects.

There are a few other verbs that take predicate adjectives or predicate nominatives. Dayley then describes various other intransitive clause and transitive clauses. He concludes the overview chapter with discussions of noun incorporation, negatives, interrogatives and imperatives.

Chapter three, Verbs and Verb Phrases, deals with all the affixes that can be attached to verbs, verbs that have different stems for plural objects, auxiliary verbs and subordinating suffixes. Next come discussions of verb derivations, such as valence changing affixes that change the voice of the verb by noun incorporation and affixes that add an instrument reference to the verb. Then there are the affixes that make passives, mediopassives, reflexives, reciprocals, absolutives and causatives.

Chapter four, Pronominals, describes all the different kinds of pronouns, including those dealing with quantifiers and numbers, time and money.

Chapter five, Nouns, Noun Phrases and Postpositions, first describes nouns in case relationships to the verb. This is followed by how nouns may be derived or compounded, and the chapter concludes with a description of noun phrases and noun phrases with postpositional markings.

Chapter six, Adjective and Adjective phrases, describes the affixes that go on adjectives, the affixes that make adjectives of verbs, verbs from adjectives, adjectives from nouns and so on. The chapter ends with comparative adjectives and modifiers that make adjective phrases.

Chapter seven, Adverbs and Particles, describes time words, place adverbs, manner adverbs, modal particles, and delimiter, intensifier, emphatic, negative, affirmative, interrogative, interjection and greeting particles.

Chapter eight, Coordination and Subordination, deals with how clauses can be combined to make complex clauses and sentences. Coordination joins two clauses together by parataxis (no conjunction) but an even intonation pattern over the whole conjoined sentence. Noun Phrases are usually coordinated by a conjunction that occurs on the end of the second constituent, often translated 'with'.

Subordinate clauses include relative clauses and nominalized (complement) clauses. Many of the relative clauses involve temporal, condition and result clauses.

Chapter nine, Phonology, is a nice presentation of the complicated phonology of the language.

Chapter ten, Texts, presents six texts. For each text, the Shoshone version is given first, then an idiomatic English version and, lastly, an interlinear word by word version.

Finally, there is a basic vocabulary list in an appendix to the book.

This grammar of Tümpisa Shoshone is comprehensive, well illustrated, and interesting. It is a worthy memorial to the language and can stand as an example for others to emulate.

JAARS Center
P.O. Box 248
Waxhaw, N.C. 28173

SHEPHERD, ALICE. 1989. *Wintu Texts*. University of California Publications in Linguistics, Volume 117. Berkeley: University of California Press. Pp. x, 497. $51.00.

Reviewed by CHARLES PECK

The Wintu people lived in Shasta, Trinity and parts of Siskiyou counties in northern California. Only a few elderly people remember the language these days.

This collection includes nearly one hundred stories and accounts. The stories are usually only a couple of pages long; a few run for several more pages. All but one text was recorded from Mrs. Grace Norton McKibbin, a remarkable woman who learned a large repertoire of stories from her grandfather and remembered them all her life. She died at the age of 92 in 1987.

The texts are all given in interlinear form followed by a free English translation. On the first text, Shepherd adds additional footnotes that explain the morphological breakdown of the complex words in the text. Otherwise, she refers the reader to the grammar by Pitkin (1984) and two dictionaries: Pitkin (1985) and Schlichter (1981).

The stories are not satisfying to a Westerner—we want stories with a problem + complication + resolution plot; these stories have a problem/activity + result plot. Sometimes the result refers to present-day situations or taboos, other times the result seems not to have any correlation with the real world.

The texts are divided into four collections: Tales (of when animals were people), Supernatural stories (in which supernatural beings are the principle actors), Ethnographic Texts and Personal Accounts. The first section, Tales, takes up over three-quarters of the book. And many of the Tales involve the Coyote as a sad-sack, stupid, silly character.

Many times, a tale makes no sense until one reads the footnote in which Mrs. McKibbin explains who the characters were. Mrs. McKibbin said she had to learn not only the tales, but also their meanings.

Many of the Tales begin with a repetitious cluster of sentences that has poetic feeling about it. This repetitiousness usually does not feel the same in the free translation. Sometimes there is also repetition at other parts of the Tales where continuous action is emphasized and when the action is an exciting part of the Tale.

Most of the sentences are simple or simple compound (with *and* or with simple juxtaposition or parataxis). A few sentences are (or could be) translated with a *but, in order to, because,* or *so that,* but the Wintu seems not to have such conjunctions. (Similar languages in Mexico have borrowed the conjunctions of Spanish to fill the holes in their own languages.)

Since the language is almost completely forgotten, this book will be of interest mostly to students of California languages and cultures. It will be a reference book of not much use to most but of invaluable use to someone who needs it.

JAARS Center
P.O. Box 248
Waxhaw, N.C. 28173

REFERENCES

Pitkin, Harvey. 1984. *Wintu grammar. University of California publications in linguistics, Volume 94*. Berkeley: University of California Press.

Pitkin, Harvey. 1985. *Wintu dictionary. University of California publications in linguistics, Volume 95*. Berkeley: University of California Press.

Schlichter, Alice. 1981. *Wintu dictionary. Reports from the survey of California and other Indian languages,* 2. Berkeley: Department of Linguistics, University of California.

CRAWFORD, JAMES M. 1989. *Cocopa dictionary. Linguistics Volume 114*. Berkeley: University of California Press. Pp. x, 534.

Reviewed by CHARLES PECK

This dictionary is the second of three volumes by Crawford. The first, *Cocopa texts,* has already been published, and the third, Cocopa grammar, is in preparation.

Cocopa is a member of the Yuman language family and is spoken in southwest Arizona and northern Mexico, by around a thousand speakers.

The dictionary begins with a brief grammar sketch—nouns take possessor inflections and verbs take inflections for person, number, aspect and tense. The grammar discussion ends with a list of all affixes. (These also appear, fortunately, in the dictionary.) Following the grammar, there is a brief chart of phonemes and some discussion of phonemes in borrowed words. This presentation is frustratingly insufficient. How would one pronounce *cxmkan?* There is mention of 'epenthetic vowels' in the list of affixes, but not in the discussion of phonemes. Nor is there any discussion of allophones—are there fricative allophones of stops?

There are approximately 5,000 entries, but perhaps only two or three thousand roots and affixes. All alternate and derived forms are listed as main entries—often from two to six entries based on the same root. Often these would be put under one main entry.

About two-thirds of the entries have no illustrative citations to indicate use. Most have cross-reference and derivation information. About a quarter have one to three illustrative sentences, so the dictionary is uneven in this respect.

There is an English-Cocopa index included. Most of the entries here refer one English word to several related Cocopa entries. This is necessary because words are listed with their derivation prefixes so that related words appear in different places.

Jaars Center
P.O. Box 248
Waxhaw, N.C. 28173

SHIPLEY, WILLIAM, ed. *In Honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics.* New York: Mouton de Gruyter, 1988. xiii + 826 pp.

Reviewed by MATTHEW S. DRYER

This volume is a collection of papers from a conference in honor of Mary Haas, one of the leading scholars in the study of American Indian languages of this century. Haas' contribution has been particularly felt through the large number of students who studied with her at the University of California at Berkeley, many of whom are represented among the authors of the papers in this collection. As with many festschrifts, many of the papers in this volume are minor, often short, papers. All deal in some way with aspects of American Indian languages, mostly languages of Canada or the United States.

Among the morphosyntactic papers, Marianne Mithun discusses various types of number marking in Central Pomo, demonstrating the inadequacies of treating such phenomena as simple mechanical agreement; she argues, as in a number of recent papers by different linguists on various languages, that what might look at first like inflectional number agreement on verbs is best viewed as a derivational process reflecting the nature of the predication rather than the number of participants in the predication. Margaret Langdon discusses extensive number suppletion in verbs in Yuman languages, found primarily with verbs of position or motion; one of her most interesting conclusions is that most of the suppletion is more recent than Proto-Yuman. Pamela Munro discusses a rather unusual phenomenon in Lakhota, Creek, Chickasaw, and Maricopa, in which verbal affixes or clitics occur which indicate a diminutive subject (or object). Noel Rude discusses the morphology associated with objects in Klamath; the system is a complex one in which the systems found on nominals, noun modifiers, and verbs, do not match up with each other. Anne Stewart argues that the switch-reference system in Conchucos Quechua cannot be described in purely syntactic terms but is sensitive to discourse topic. Michelle Emanation defends the appropriateness of applying the term "passive" to what has also been called an inverse construction in Nootka. Eloise Jelinek and Fernando Escalante discuss an unusual possessive construction in Yaqui in which the possessed noun appears as main predicate with verbal morphology while its modifiers exhibit object inflection. Amy Dahlstrom gives a detailed description of the use of independent pronouns in Fox, going well beyond the oft-used notion that independent pronouns in head-marking languages are "used for emphasis."

While all of the papers mentioned in the preceding paragraph are of typological and theoretical importance, the contribution by Johanna Nichols differs in that it is a cross-linguistic study, which, while concentrating on the languages of Canada and the United States, discusses languages from various parts of the world. Although it is one of the longer papers in the volume, it is packed with a large number of insightful observations about the formal properties associated with the distinction between alienable and inalienable possession. Among her conclusions is that the distinction is ultimately a lexical one, with the exact set of lexical items exhibiting inalienable possession generally defying exact semantic characterization. Inalienable possession is often restricted to kin terms and/or body parts, but not always. The distinction is related to her own distinction between head-marking and dependent-marking languages: inalienable possession is more often associated with head-marking than alienable possession is. In fact the distinction rarely involves two types of dependent marking. Alienable possession is often more marked morphologically than inalienable possession, either in terms of the number of morphemes or in terms of the number of words.

If my comments on the remaining papers in the volume are briefer, this reflects in part my own personal interests. Nine of the papers deal with phonology or morphophonology, James Armagost on Comanche, Robert Oswalt on Kashaya, Frank Siebert on Penobscot, Charles Ulrich on Choctaw, David Cline on Oklahoma Seminole, Heather Hardy and Timothy Montler on Alabama, Geoffrey Kimball on Koasati, and separate papers by Jeff Burnham and Larry Hagberg on Mayo. Eight papers deal with comparative-historical questions: Catherine Callaghan on Utian, Brent Galloway on Central Salish, Dale Kinkade on Salishan, James Crawford on Timucua, Janine Scancarelli on Cherokee, Ives Goddard on Cheyenne, Paul Proulx on general issues with reference to Algic, and Scott Delancey, Carol Genetti, and Noel Rude with the most detailed published evidence for a genetic relationship between Sahaptian, Klamath, and Tsimshianic, three groups often hypothesized to be related to the Penutian languages of California. Three papers deal with areal-contact questions: Robert Rankin on Quapaw, David Shaul on Esselen, and Eugene Casad on Cora. Two papers discuss numeral systems: Pamela Bunte and Robert Franklin on San Juan Southern Paiute and Katherine Turner on Salinan. Two papers discuss issues of style: Sheldon Klein on Kawaiisu and William Shipley on Maidu. Wick Miller discusses the semantics of positional words in Guarijío. Emanuel Drechsel discusses resemblances between the thinking of Edward Sapir and that of Wilhelm von Humboldt, not to be confused with his better-known younger brother, Alexander von Humboldt. The final paper by Karl Teeter is an anecdotal account of some of the practical and human difficulties in completing a lexicon.

Through her supervisory role at Berkeley, especially via the Survey of California Indian languages, Mary Haas has been indirectly responsible for a considerable amount of the descriptive work on American Indian languages of Canada and the United States over the past forty years. Sadly, because of the critical state of many underdescribed languages and because of current fashions in the field, the need for somebody like Mary Haas to continue that role in the field is now greater than ever.

Dept. of Linguistics
SUNY at Buffalo
Buffalo, NY 14260

STIG JOHANSSON and KNUT HOFLAND. *Frequency Analysis of English Vocabulary and Grammar Based on the LOB Corpus.* Volume 2: *Tag Combinations and Word Combinations.* Oxford: Clarendon Press, 1989. Pp. [vi] + 380.

Reviewed by JOHN ALGEO

The LOB (Lancaster-Oslo/Bergen) Corpus is a collection of texts of British English which were originally printed in 1961. Totaling a million words, the Corpus consists of 500 samples containing approximately 2000 words each, diversified among 15 categories of writing to match the similar Brown Corpus of American English. The corpus is available in computerized form with each word and punctuation mark accompanied by a tag. Most of the 154 tags (listed in an appendix of the work under review, 377–380) identify grammatical classes of words.

Volume 1 of *Frequency Analysis* includes an introduction briefly describing the corpus, the tagging procedure, and the tag set. The body of that volume reports the frequency of each tag in each of the 15 categories of text, collectively in the nine categories of informative writing and in the six of imaginative writing, and in the corpus as a whole. Most significantly, it reports the frequency of each word in the corpus and of every tag for each word. For example, the word *round* occurs 335 times in the corpus, 148 times as a preposition, 29 times as an adjective, 19 times as a singular count noun, 138 times as an adverbial particle, and once as the base form of a verb.

Volume 2 continues the statistical analysis on a more sophisticated level, reporting the frequencies of tag and word combinations. Pages 15–93 list all the tags, together with the tags that precede and follow them. For each combination of tags is given (1) the frequency of that combination, (2) the percentage of the combination out of the total number of instances of the focused tag, and (3) the percentage of the combination out of the total number of instances of the accompanying tag.

For example, the tag IN, identifying prepositions, is preceded by a singular count noun (NN) 42,923 times, which is 34.77 percent of the instances of prepositions and 28.85 percent of the instances of singular count nouns. Prepositions are followed by the kind of article that can be either singular or plural (such as *the* and *no*) 35,500 times, which is 28.76 percent of the instances of prepositions and 50.56 percent of the instances of such articles. From the tag combinations we can learn a good bit about the sequences in which grammatical classes tend to occur. The fact that articles like *the* follow a preposition half the time is not otherwise obvious.

Tag combinations occurring fewer than ten times are reported collectively, but not specified individually. So prepositions are preceded 118 times (.09 percent of their occurrences) by some tag other than the 90 for which statistics are individually reported. They are followed 120 times (.1 percent) by some tag other than the 82 individually reported ones.

Most of the volume, pages 95–376, lists combinations of frequently occurring verbs, adjectives, nouns, and adverbial particles with noteworthy collocations. For verbs and adjectives, collocations with the following tags and with particular words are both given. For nouns and adverbial particles, only lexical collocations are given.

It is instructive to compare the frequencies in the LOB Corpus with those in the computerized files of the Survey of English Usage at University College London, a million-word corpus that includes both written and spoken texts (Quirk 1968), and with the Brown Corpus. Thus the verb *accept,* in its various inflected forms, occurs in LOB a total of 247 times, of which 82 instances are followed by a determiner (as in *accept the responsibility*). It occurs 193 times in Brown (Francis and Kucera 1982), and 207 times in the Survey.

There are two characteristically British collocations for *accept.* In LOB, 9 instances of *accept* are followed by the subordinating conjunction *that* (*accept that* something is the case) and 2 instances by the infinitive marker (*accept to* do something). The Survey corpus includes 15 instances of *accept that,* of which the following are typical:

> They are prepared to accept that the white man in Rhodesia is expendable. (w12.4.88)
>
> Will [undecipherable noise] Secretary of State [uh] accept that I welcome his plans to consult with what he calls interested parties? (s11.4.246–248)

A large number of the spoken examples are from parliamentary language. Benson, Benson, and Ilson (1986) label the construction, which they illustrate by "accept that the proposal may be defeated," as formal and British English. Quirk et al. (1985, 16.31) list *accept* as one of the verbs taking *that*-clauses as complement but specify no usage limitation.

There are no instances in the computerized Survey corpus of *accept to.* Paul Beale (personal communication) has pointed to a use from 1962, a year after those of the LOB corpus: "Before I had had time to wipe the cream of my third chocolate éclair from my lips, I discovered that I had accepted to become his business manager,'' as well as a current one: ''He [Saddam Hussein] has accepted to move out [from Kuwait].'' Sylvia Chalker (personal communication) points out that the construction is listed as a new word by Ayto (1990) with a citation from 1988. It is a new form in that it is not listed in dictionaries. The LOB Corpus, however, has the oldest attested use of the form.

Some useful information had to be omitted from *Frequency Analysis* for lack of space. The LOB corpus includes 9 instances of *cater* as a verb, which are not enough to get the verb into the collocational lists of volume 2. The rule for inclusion is that at least one inflected form must occur more than ten times. Some such arbitrary limitation was necessary, but it unfortunately leaves interesting information unavailable—in this case, the prepositional complement after *cater.* British English prefers *cater for,* and American *cater to.* The Survey files include 12 instances of the verb *cater,* 11 with *for* as a complement, 1 without a complement, and none with *to.* It would be useful to have such information from LOB readily available.

The adjective *old,* not surprisingly, is usually followed by a noun. But the list of particular nouns shows major collocations (highest frequency first) to be *old man/men, old house/houses, old friend/friends, old woman/women, old age, old people.* The difference in collocation for the adjectives *British* and *American* is perhaps indicative of differences in the British view of themselves and us. Most frequent collocations for the former are *British government, British Museum, British public, British Isles, British commonwealth, British people.* The most frequent collocations for the latter are *American army, American tanks, American revolution.*

The noun list emphasizes preceding adjectives and verbs. For *office,* the most frequent adjective combinations are *foreign office* and *colonial office,* neither of which would be expected in an American list. Preceding verbs are more international: *take office, hold office, left the office.*

The list of adverbial particles gives their collocation with verbs, thus showing the number of such combinations in English. *Out* is the most frequent of these particles in the LOB Corpus; *up* is the second most frequent. The latter has the combinations *add up, bind up, blow up, break up, bring up, build up, buy up, call up, catch up, check up, clean up, clear up, climb up, come up, cover up, cut up, divide up, draw up, dress up, drive up, eat up, end up, face up, fill up, finish up, follow up, gather up, get up, give up, glance up, go up, grow up, hang* *up, heat up, hold up, hurry up, jump up, keep up, kick up, lay up, lead up, lift up, light up, link up, live up, lock up, look up, make up, measure up, meet up, mix up, move up, open up, pack up, pay up, pick up, pull up, put up, reach up, ride up, ring up, roll up, run up, save up, sell up, set up, shoot up, show up, shut up, sit up, snap up, speak up, split up, spring up, stand up, start up, step up, stick up, stir up, straighten up, strike up, swallow up, take up, think up, throw up, tie up, turn up, use up, wake up, walk up, work up, wrap up, write up, yield up*—and that is not counting multiple senses of many of those forms or extended forms like *go on up* or *stand up to.*

Characteristically British use can be seen in the particles too, for example, the choice between *around* and *round.* Combinations that occur with both are the following (with their frequency in parentheses); *go around* (5), *round* (10); *look around* (7), *round* (12); *turn around* (3), *round* (11). The total frequency of the particles in the LOB Corpus are *around* 86, *round* 138; in the Brown Corpus, *around* (240), *round* (14). Thus, while British strongly prefers *round,* American overwhelming prefers *around.*

Despite the inevitable limitations on how much information could be included in this volume, it is an extremely useful work. Its usefulness will be even greater when it is supplemented by other works giving comparable statistics about tag and word combinations for the Brown Corpus and other corpuses like those of the Survey of English Usage and the International Corpus of English, now in preparation under the general direction of Sidney Greenbaum (Greenbaum 1990a, 1990b, 1991). For the present, however, this volume is our best and most available source of information about the statistics of English collocations.

English Department
University of Georgia
Athens, GA 30602

REFERENCES

Ayto, John. 1990. *The Longman register of new words: Volume Two.* Harlow, Essex: Longman.

Benson, Morton, Evelyn Benson, and Robert Ilson. 1986. *The BBI combinatory dictionary of English: A guide to word combinations.* Amsterdam: Benjamins.

Francis, W. Nelson, and Henry Kucera. 1982. *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Greenbaum, Sidney. 1990a. "Standard English and the International Corpus of English." *World Englishes* 9:79–83.

———. 1990b. "International Corpus of English: A progress report." *World Englishes* 9:121.

———. 1991. "The development of the International Corpus of English." In *English corpus linguistics,* ed. K. Aijmer and B. Altenberg. London: Longman.

Quirk, Randolph. 1968. "The Survey of English Usage." In *Essays on the English language: Medieval and modern.* London: Longman.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A comprehensive grammar of the English language.* London: Longman.